PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C12N 7/00, 7/01, 15/00, 15/11, 15/86, A61K 39/00, 39/29

A1

(11) International Publication Number: WO 94/12617

(43) International Publication Date: 9 June 1994 (09.06.94)

(21) International Application Number: PCT/US93/11474

(22) International Filing Date: 24 November 1993 (24.11.93)

(30) Priority Data:
982,211 25 November 1992 (25.11.92) US

(71) Applicant: INTERNATIONAL BIOTECHNOLOGY LABORATORIES, INC. [US/US]; 67 Rogers Street, Cambridge, MA 02142 (US).

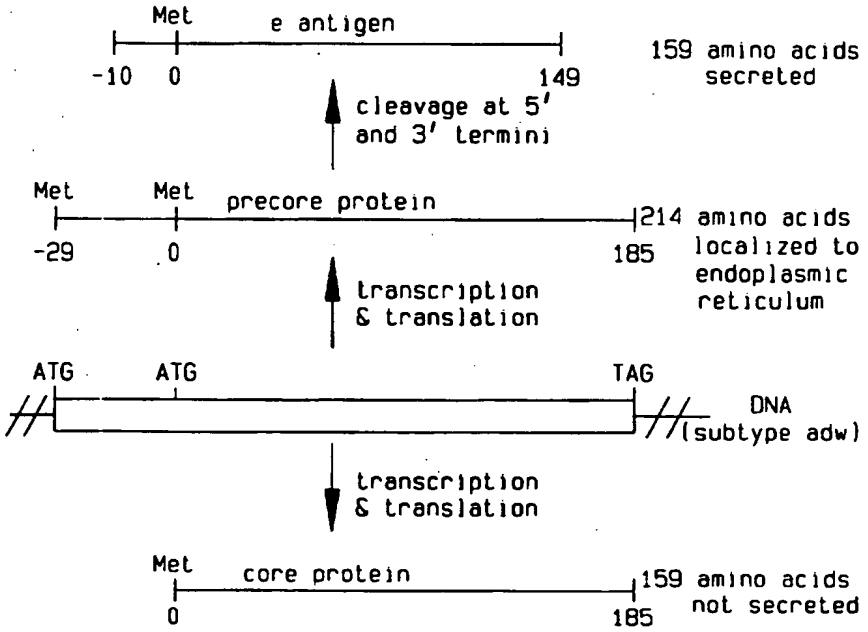
(72) Inventors: SOUW, Peter, T., S.; 456 Belmont Street, #12, Watertown, MA 02172 (US). O'KEEFE, Rhonda, Wilson; 57 Glen Drive, Hudson, NH 03051 (US). LEWIS, Tatyana; 49 Glendale Road, Newton, MA 02159 (US). BERNSTINE, Edward, G.; 376 Commercial Street, Boston, MA 02109 (US).

(74) Agents: MISROCK, S., Leslie et al.; Pennie & Edmonds, 1155 Avenue of the Americas, New York, NY 10036 (US).

(81) Designated States: AU, BB, BG, BR, BY, CA, CZ, FI, HU, JP, KR, KZ, LK, LV, MG, MN, MW, NO, NZ, PL, RO, RU, SD, SK, UA, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: HEPATITIS B VIRUS VACCINES



(57) Abstract

The present invention provides recombinant vaccinia viruses which are capable of expressing a plurality of immunogenic HBV epitopes comprising at least one core antigen epitope, at least one surface antigen epitope, or any combination of the foregoing. Preferred viruses containing plural HBV antigens include but are not limited to a combination of one of S, MS, LS, and full-length core, coreΔ8 (a core deletion mutant), preS1-coreΔ8, preS1-complete core, core-preS1*, core-preS2, or core-S* (the asterisks denoting the presence of an immunogenic fragment of the preceding antigen; hyphens denoting a fusion protein). As shown in the figure, the core and/or surface antigen epitopes can be expressed by the same recombinant virus, under the control of different promoters active in vaccinia virus.

HEPATITIS B VIRUS VACCINES

1. INTRODUCTION

The present invention relates to recombinant vaccinia viruses which are capable of expressing a plurality of hepatitis B virus (HBV) epitopes comprising core antigen epitopes, surface antigen epitopes, or any combination of the foregoing. The vaccinia viruses are useful in vaccine formulations to prevent or treat hepatitis or other undesirable consequences of HBV infection.

2. BACKGROUND OF THE INVENTION

2.1. HEPATITIS B VIRUS

Hepadnaviruses, which include human hepatitis B virus (HBV) (Barker et al., 1975, Am. J. Med. Sci. 270:189-196), are hepatropic, lead to persistent infection, and have a common structure. Large concentrations of hepadnaviruses have been detected in the blood of infected organisms. Hepadnavirus virions are approximately 42 nm in diameter and have a double-walled structure consisting of an envelope and nucleocapsid (reviewed in Ganem and Varmus, 1987, Ann. Rev. Biochem. 56:651-693 and Tiollais et al., 1985, Nature (London) 317:489-495). The outer coat, or envelope, contains three surface antigens, designated major or small (S), middle (MS), and large (LS), as well as carbohydrates and lipids. Hepatitis B surface antigen (HBsag) is found on all three of these surface antigen proteins. The nucleocapsid, or inner coat, contains a circular DNA (3.0-3.3 kb in length), a DNA polymerase, protein kinase activity, and hepatitis B core antigen (HBcAg).

The hepadnavirus genome is a small, circular, partly double stranded DNA molecule (reviewed in Ganem and Varmus, *supra*, and Tiollais et al., *supra*). The minus strand is linear and of a fixed length, 3-3.2 kb. In contrast, the plus strand is of variable length, ranging from 50-100% of that of the minus strand.

The results from analyses of nucleotide sequences of the cloned HBV genome indicated that the minus strand contains four major open reading frames (ORFs), S, C, P, and X (reviewed in Ganem and Varmus, *supra*, and Tiollais, *supra*). ORF S encodes a total of 389-400 amino acids, depending upon subtype. ORF S is divided into the preS1 region (encoding 108-119 amino acids), preS2 region (encoding 55 amino acids), and S gene (encoding 226 amino acids) in 5′ to 3′ order, each of which begins with an ATG codon capable of functioning as a translation initiation site *in vivo*. Gene S encodes small (S) antigen, the HBsAg, which is the major envelope protein. A 500 nucleotide sequence upstream of the S gene contains the preS1 and preS2 sequences. Middle (MS) protein is encoded by the preS2 and S regions, and large (LS) protein is encoded by the preS1, preS2 and S regions. ORF C, which codes for HBcAg, is divided into the C gene and pre-C region. ORF P encodes the viral polymerase. ORF X encodes a protein which enhances expression of the other viral genes.

Core-related antigen production *in vivo* is a complex process involving two different precursors, precore and core, and post translational truncation at both amino and carboxy termini (Fig. 1). Two main forms of core-related antigens result from this processing: e antigen, which is made from precore and found mainly extracellularly in infected individuals,

and core, which is found almost exclusively intracellularly, usually within the nucleus of the cell. Putative nuclear localization sequences of the hepatitis B core protein include a set of direct repeats having the amino acid sequence PRRRRSQS (SEQ ID NO:1) located in tandem in the protamine-like domain of core. Yeh et al. (1990, J. Virol. 64:6141-6147) suggest that these repeats are necessary for nuclear localization. Eckhardt et al. (1991, J. Virol. 65:575-582) suggest that other sequences may also be involved.

The structure and expression of the hepatitis B viral genome has been studied by recombinant DNA techniques. U.K. Patent Publication No. GB 2034323A (published June 4, 1980) discloses the cloning and expression of HBV DNA. Moriarty et al. (1981, Proc. Natl. Acad. Sci. USA 78:2606-2610) describes the construction and expression of a recombinant simian virus containing a fragment of HBV DNA encoding the HBsAg; 22 nm surface antigen particles were shown to be excreted into the cell culture medium. European Patent Publication No. 0020251 (published December 10, 1980) describes the construction of and expression from recombinant expression vectors encoding HBV proteins. European Patent Publication No. 0013828 (published August 6, 1980) discloses recombinant DNA vectors which contain fragments of HBV DNA, isolated from Dane particles, which encode proteins with HBV antigenicity.

2.2. BIOLOGY OF HBV INFECTION

The liver disease called hepatitis B is a widespread and serious health problem resulting from infection by the hepatitis B virus. A large proportion of people living in regions of poor medical

care are chronically infected with the virus and also face an elevated risk of acquiring liver cancer, since HBV is associated with liver carcinoma. HBV is endemic in Asia, and causes one of the largest health problems in that region, where about 10% of the population are chronically infected carriers. The toll of this figure on the health of Asians is severe; in addition to the impact of acute HBV infection, chronic carriers are at vastly increased risk from cirrhosis and primary hepatocellular carcinoma (PHC). PHC is the second most prevalent cancer in Asia, accounting for up to 22% of cancer deaths in some countries.

HBV infection has been observed to be highly polymorphic, ranging from inapparent forms in which individuals experience mild or no liver injury to acute hepatitis B, a moderately severe illness characterized by hepatocellular injury and inflammation to severe chronic liver disease (reviewed in Ganem and Varmus, 1987, Ann. Rev. Biochem. 56:651-693). People infected with HBV are often unaware of it. After a two- to six-month incubation period, HBV infection can lead to acute hepatitis and liver damage, which cause abdominal pain, jaundice, elevated levels of certain enzymes in the blood and other symptoms. At this stage, hepatitis B viral infection can be diagnosed by detecting hepatitis B surface antigen (HBsAg) in a patient's blood serum. More frequently, the disease remains permanently asymptomatic. Although many chronic carriers appear healthy, they can still transmit the hepatitis B virus to those with whom they have close contact, thereby starting the cycle of disease anew.

Variation in host immune response to virus infected cells appears to be a major determinant in

the degree of severity of liver damage in individuals. Both humoral and cellular immune responses to HBcAg and HBsAg have generally been observed during acute and chronic HBV infection (reviewed in Robinson, 1986, *Fundamental Virology*, Fields, B.N. and Knipe, D.M. (eds.), Raven Press, N.Y., pp. 657-679).

In addition to causing acute and chronic liver disease, results from epidemiological and molecular biological studies indicate a connection between hepadnavirus infection and hepatoma (liver cancer) (reviewed in Ganem and Varmus, *supra*; Machowiak, 1987, Am. J. Med. 82:79-97; and Tiollais et al., 1985, Nature (London) 317:489-495. There is a strong correlation between high rates of chronic HBV infection and incidence of hepatoma, *e.g.*, in Southeast Asia and equatorial Africa.

Usually a patient with acute hepatitis will recover completely from the disease. After recovery, the patient continues to produce low levels of antibodies specific for HBV, and thus is immune to HBV for several years. In the event of a new HBV attack, the antibody level rises quickly and neutralizes the virus. In chronically-infected patients, the virus survives in the liver, and may cause cirrhosis and, eventually, hepatocarcinoma. Chronic infection occurs in approximately 5-10% of infected adults and 80% of infected infants. Hepatitis B is, therefore, primarily a disease of infants in developing countries, and is mostly confined to adults in Western countries.

2.3. VACCINATION AND THERAPIES FOR HBV INFECTION

Current methods used to treat HBV infection may be divided into two categories, antivirals and immunomodulatory agents (reviewed in Hoofnagle, Ann.

Int. Med. 107:414-415). One analog that has been widely tested is vidarabine (adenine arabinoside), an adenine analog with potent activity against herpesviruses. While the inhibition of viral replication was observed to be nearly complete and sustained in some patients, the inhibition of replication was partial and transient in other patients. One drawback involved in using this drug is that it is highly water insoluble and has to be administered intravenously as a constant infusion. A monophosphorylated derivative of vidarabine, vidarabine phosphate, can be administered by rapid intravenous infusion or intramuscularly. Even though vidarabine phosphate therapy resulted in the clearance of serum hepatitis B virus DNA, it did not lead to a sustained improvement in the accompanying liver disease. Other antivirals that have been tested for their effectiveness in treating HBV infection include acyclovir and suramin.

Studies have also been undertaken in which patients were treated with alpha-interferons for 1-6 months. It was found that only 25-40% of the patients responded to this therapy. The combination of vidarabine phosphate and human leukocyte interferon proved to be toxic. Other therapies that have been studied include the administration of interleukin-2, gamma interferon, and short course corticosteroids.

U.S. Patent No. 4,471,901 discloses a vaccine comprising a 22 nm polypeptide particle made of mature hepatitis B surface antigen.

Szmuness et al. (1981, J. Med. Virol. 8:123) report that 2-4% of the general population are nonresponsive to the existing recombinant hepatitis vaccine, which contains only the S antigen. Milich et al. (1986, J. Immunol. 137:315) suggest that this lack

of an anti-S response may be overcome by immunizing with preS epitopes in combination with S. Neurath et al. (EPO publication no. 243,913) disclose an HBV vaccine which includes amino acid sequences from within the preS region of the HBV surface antigen linked to a lipid vesicle. Other groups disclose various surface antigen constructs for vaccine use (U.S. Patent No. 4,816,564; U.S. Patent No. 4,683,136; PCT International Publication No. WO 8810301). Immunization of chimpanzees with recombinant or purified HBcAg has been shown to have some protective effect against HBV challenge (Tabor and Geretz, 1984, Lancet i:172; Murray et al., 1984, EMBO J. 3:645). However, vaccines employing purified HBV antigens are disadvantageous due to the laborious purification procedures required.

A number of groups have made constructs consisting of fusions to the core antigen expressed in bacteria, yeast or vaccinia (see, *e.g.*, U.S. Patent No. 4,859,465; U.S. Patent Nos. 4,818,527 and 4,882,145). In some cases core particles are formed which expose the foreign antigenic determinants on their surface (Clarke et al., 1987, Nature 330:381-384; Borisova et al., FEBS Letters 259:121-124).

Potential live vaccines consisting of vaccinia virus recombinants expressing HBV surface antigens have been reported (Smith et al., 1983, Nature 302:490-495; Cheng and Moss, 1987, J. Virol. 61(4):1286-1290; Cheng et al., 1986, J. Virol. 60(2):337-344; Paoletti et al., 1984, Proc. Natl. Acad. Sci. USA 81:193-197; Moss et al., 1984, Nature 311:67-69; U.S. patent application Serial No. 6-445,892 filed December 1, 1982).

2.4. VACCINIA VIRUS

Vaccinia virus is a poxvirus (for a review, see Moss, 1990, *Virology*, 2d ed., Ch. 74, Fields et al., eds., Raven Press, Ltd., New York, pp. 2079-2111) which replicates within the cytoplasm of infected cells. Vaccinia virus contains a linear double-stranded genome of approximately 187 kilobase pairs, encoding approximately 200 proteins. A large number of proteins are encoded by the virus because the virus establishes infection in the cytoplasm of the cell rather than in the nucleus, and therefore must provide its own set of enzymes to replicate its DNA and to transcribe its genes.

A vaccine approach has been described, involving the use of vaccinia virus as a vector to express foreign genes inserted into its genome (Mackett et al., 1982, Proc. Natl. Acad. Sci. USA 79:7415-7419; Mackett et al., 1984, J. Virol. 49:857-864; Panicali and Paoletti, 1982, Proc. Natl. Acad. Sci. USA 79:4927-4931; PCT International Application No. WO 91/12318; European Patent Publication No. 83,286; Murphy, 1989, Res. Virol. 140:463-491; Moss and Flexner, 1987, Ann. Rev. Immunol. 5:305-324; Hruby, 1990, Clin. Microbiol. Rev. 3(2):153-170; U.S. Patent No. 4,603,112; U.S. Patent No. 4,722,848; U.S. patent application Serial No. 7-072,455 filed July 13, 1987). Upon introduction into host animals, the recombinant vaccinia virus expresses the inserted foreign gene and may thereby elicit a host immune response to such gene products.

3. SUMMARY OF THE INVENTION

The present invention provides recombinant vaccinia viruses which are capable of expressing a plurality of immunogenic HBV epitopes comprising at

least one core antigen epitope, at least one surface antigen epitope, or any combination of the foregoing. In one embodiment, a plurality of core and/or surface antigen epitopes is expressed. Core antigen epitopes are selected from the group consisting of wild-type core antigen, e antigen, and derivatives thereof. Surface antigen epitopes are those encoded by ORF S, and are selected from the group consisting of S, MS, LS, preS2, preS1, and derivatives thereof. "Epitope" as used herein refers to an antigenic determinant, *i.e.*, an amino acid sequence that is capable of being immunospecifically bound by an antibody. Thus, for example, a protein containing a core antigen epitope can be detected by observing the ability of such protein to be bound by an anti-core protein antibody. The core antigen derivatives provided by the invention are those derivatives (including but not limited to fragments) which display core antigen antigenicity, *i.e.*, are capable of being bound by an anti-core antibody. Similarly, the surface antigen derivatives provided by the invention are those derivatives (including but not limited to fragments) which display surface antigen antigenicity, *i.e.*, are capable of being bound by an anti-surface antigen antibody.

In a specific embodiment, a combination of the foregoing core and/or surface antigen epitopes are expressed as a fusion protein in the vaccinia viruses of the invention. "Fusion protein," as used herein, refers to a protein comprising an amino acid sequence from a first protein covalently linked via a peptide bond at its carboxy terminus to the amino terminus of an amino acid sequence from a second, different protein.

In one embodiment, a vaccine formulation of the invention contains a single type of recombinant

vaccinia virus of the invention. In another embodiment, a vaccine formulation comprises a mixture of two or more recombinant viruses of the invention. Preferred viruses containing plural HBV antigens include but are not limited to a combination of one of S, MS, LS, and full-length core, coreΔ8 (see Section 6.1.2 *infra*), preS1-coreΔ8, preS1-complete core, core-preS1*, core-preS2, or core-S* (the asterisks denoting the presence of an immunogenic fragment of the preceding antigen; hyphens denoting a fusion protein).

In various embodiments, the core and/or surface antigen epitopes can be expressed by the same recombinant virus, under the control of different promoters active in vaccinia virus. In a preferred aspect, two divergently oriented promoters are used. In a specific embodiment, a recombinant vaccinia virus expresses a surface-core (written in the amino to carboxy-terminal order) or core-surface fusion protein (*e.g.*, core-preS1, core-preS2, core-S, or preS1-core) under the control of a first promoter, and expresses a surface antigen (*e.g.*, MS) under the control of a second promoter.

In another specific embodiment, a mixture of recombinant vaccinia viruses is provided. In one embodiment, such a mixture comprises a first virus capable of expressing core-preS1, a second virus capable of expressing core-preS2, and a third virus capable of expressing core-S or core-S*. In another embodiment, such a mixture comprises a first virus capable of expressing core-preS1 or core-preS1*, a second virus capable of expressing core-preS2, and a third virus capable of expressing S. In another embodiment, such a mixture comprises a first virus capable of expressing core-preS1 or core-preS1*, and a

second virus capable of expressing MS. In yet another embodiment, such a mixture comprises a first virus capable of expressing core-preS2, and a second virus capable of expressing S.

4. DESCRIPTION OF THE FIGURES

Figure 1. Schematic diagram of HBV precore and core processing to produce core antigen and e antigen.

Figure 2. Nucleotide sequence (SEQ ID NO:2) of the p7.5 promoter.

Figure 3. Schematic diagram of plasmid pGS53.

Figure 4. Nucleotide sequence (SEQ ID NO:3) of the p11 promoter.

Figure 5. Schematic diagram of plasmid pSC10.

Figure 6. Nucleotide sequence (SEQ ID NO:4) of the modified p7.5 promoter.

Figure 7. Schematic diagram of plasmid pSC59.

Figure 8. Schematic diagram of the HBV genome, illustrating the core and surface antigen-encoding regions.

Figure 9. Schematic diagram of plasmid pAM6.

Figure 10. Schematic diagram of plasmid pT7T318.

Figure 11. Schematic diagram of plasmid pRO-01. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 12. Schematic diagram of plasmid pRO-02. A restriction enzyme site shown in

parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 13. Schematic diagram of plasmid pRO-03.

Figure 14. Schematic diagram of plasmid pRO-05. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 15. Schematic diagram of plasmid pRO-06. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 16. Schematic diagram of plasmid pT7T3/S. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 17. Schematic diagram of plasmid pT7T3/S2.

Figure 18. Schematic diagram of plasmid pLEH-01.

Figure 19. Schematic diagram of plasmid pLEH-02. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 20. Schematic diagram of plasmid pLEH-03. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 21. Schematic diagram of plasmid pLEH-04. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 22. Nucleotide sequence (SEQ ID NO:5) and amino acid sequence (SEQ ID NO:6) of LS.

Figure 23. Nucleotide sequence (SEQ ID NO:7) and amino acid sequence (SEQ ID NO:8) of MS.

Figure 24. Nucleotide sequence (SEQ ID NO:9) and amino acid sequence (SEQ ID NO:10) of S.

Figure 25. Schematic diagram of plasmid pRO-09B. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 26. Schematic diagram of plasmid pT7T3/C. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 27. Schematic diagram of plasmid pT7T3/CHB.

Figure 28. Schematic diagram of plasmid pT7T3/CRB.

Figure 29. Schematic diagram of plasmid pT7T3/CODM.

Figure 30. Nucleotide sequence (SEQ ID NO:12) and amino acid sequence (SEQ ID NO:13) of coreΔ8.

Figure 31. Schematic diagram of plasmid pCRB/24. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 32. Schematic diagram of plasmid pCODM/24.

Figure 33. Nucleotide sequence (SEQ ID NO:14) and amino acid sequence (SEQ ID NO:15) of full-length core in plasmid pCODM/24.

Figure 34. Schematic diagram of plasmid pGS53X.

Figure 35. Schematic diagram of plasmid pGS53/S.

Figure 36. Schematic diagram of plasmid pGS53/S2.

Figure 37. Schematic diagram of plasmid pGS53/S1.

Figure 38. Schematic diagram of plasmid pHTL-8.

Figure 39. Schematic diagram of plasmid pHTL-9.

Figure 40. Schematic diagram of plasmid pHTL-10.

Figure 41. Schematic diagram of plasmid pHTL-10M.

Figure 42. Schematic diagram of plasmid pHTL-25. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 43. Schematic diagram of plasmid pHTL-26.

Figure 44. Nucleotide sequence (SEQ ID NO:25) and amino acid sequence (SEQ ID NO:26) of core-preS1*.

Figure 45. Schematic diagram of plasmid pHTL-27.

Figure 46. Nucleotide sequence (SEQ ID NO:29) and amino acid sequence (SEQ ID NO:30) of core-preS2.

Figure 47. Schematic diagram of plasmid pHTL-28.

Figure 48. Nucleotide sequence (SEQ ID NO:33) and amino acid sequence (SEQ ID NO:34) of core-S*.

Figure 49. Schematic diagram of plasmid pSC10ΔlacZ.

Figure 50. Schematic diagram of plasmid pDPV. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 51. Schematic diagram of plasmid pDPV-01.

Figure 52. Schematic diagram of plasmid pRO-10.

Figure 53. Schematic diagram of plasmid pRO-16. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 54. Schematic diagram of plasmid pRO-11. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 55. Schematic diagram of plasmid pRO-17. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 56. Schematic diagram of plasmid pRO-22. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 57. Schematic diagram of plasmid pRO-12.

Figure 58. Schematic diagram of plasmid pRO-13B. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 59. Schematic diagram of plasmid pRO-18. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 60. Schematic diagram of plasmid pRO-19.

Figure 61. Schematic diagram of plasmid pRO-19/24. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 62. Schematic diagram of plasmid pHTL-5. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 63. Schematic diagram of plasmid pHTL-6.

Figure 64. Schematic diagram of plasmid pHTL-7.

Figure 65. Schematic diagram of plasmid pHTL-11.

Figure 66. Schematic diagram of plasmid pHTL-12.

Figure 67. Schematic diagram of plasmid pRO-16M. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 68. Schematic diagram of plasmid pHTL-30.

Figure 69. Schematic diagram of plasmid pHTL-13.

Figure 70. Schematic diagram of plasmid pHTL-15. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 71. Schematic diagram of plasmid pHTL-17. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 72. Schematic diagram of plasmid pHTL-33.

Figure 73. Schematic diagram of plasmid pHTL-34.

Figure 74. Schematic diagram of plasmid pHTL-35.

Figure 75. Schematic diagram of plasmid pHTL-36.

Figure 76. Schematic diagram of plasmid pPB-05. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 77. Schematic diagram of plasmid pPB-09.

Figure 78. Schematic diagram of plasmid pET-3d. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 79. Schematic diagram of plasmid pT7/core. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 80. Schematic diagram of plasmid pPB-04.

Figure 81. Nucleotide sequence (SEQ ID NO:44) and amino acid sequence (SEQ ID NO:45) of preS1-coreΔ8.

Figure 82. Schematic diagram of plasmid pRO-21. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 83. Nucleotide sequence (SEQ ID NO:48) and amino acid sequence (SEQ ID NO:49) of preS1-full-length core.

Figure 84. Schematic diagram of plasmid pRO-23. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 85. Schematic diagram of plasmid pHTL-23. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 86. Schematic diagram of plasmid pHTL-24. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 87. Schematic diagram of plasmid pHTL-18. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 88. Schematic diagram of plasmid pHTL-14. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 89. Schematic diagram of plasmid pHTL-31.

Figure 90. Schematic diagram of plasmid pHTL-32. A restriction enzyme site shown in parentheses means that the site was present in the parental plasmid construct but was destroyed during cloning procedures.

Figure 91. A representative Western blot of polypeptides in cells infected by recombinant vaccinia viruses, probed with anti-S antibodies. Lane A contains proteins isolated from nonrecombinant (wild-type) vaccinia virus. Lane B contains proteins from a recombinant virus which does not express any S-related antigens. Lanes C (pRO-18) and D (pGS53/S1) are samples from recombinant viruses expressing the two forms of the LS antigen, nonglycosylated and glycosylated polypeptides of expected molecular weights 39 kD and 42 kD. Lanes E (pGS53/S) and F (pHTL-8) contain proteins from recombinant viruses expressing the two forms of the S antigen, with

nonglycosylated and glycosylated forms having expected molecular weights of 24 kD and 27 kD. Lanes G (pHTL-14), H (pGS53/S2) and I (pRO-23) contain proteins from recombinant viruses expressing the two forms of the MS antigen, with singly and doubly glycosylated forms of expected molecular weight 33 kD and 36 kD, respectively.

5. DETAILED DESCRIPTION OF THE INVENTION

The present invention provides recombinant vaccinia viruses which are capable of expressing a plurality of immunogenic HBV epitopes comprising at least one core antigen epitope, at least one surface antigen epitope, or any combination of the foregoing. In one embodiment, a plurality of core and/or surface antigen epitopes is expressed. Core antigen epitopes are selected from the group consisting of wild-type core antigen, e antigen, and derivatives thereof. Surface antigen epitopes are those encoded by ORF S, and are selected from the group consisting of S, MS, LS, preS2, preS1, and derivatives thereof. "Epitope" as used herein refers to an antigenic determinant, *i.e.*, an amino acid sequence that is capable of being immunospecifically bound by an antibody. Thus, for example, a protein containing a core antigen epitope can be detected by observing the ability of such protein to be bound by an anti-core protein antibody. The core antigen derivatives provided by the invention are those derivatives (including but not limited to fragments) which display core antigen antigenicity, *i.e.*, are capable of being bound by an anti-core antibody. Similarly, the surface antigen derivatives provided by the invention are those derivatives (including but not limited to fragments) which display

surface antigen antigenicity, *i.e.*, are capable of being bound by an anti-surface antigen antibody.

In a specific embodiment, a combination of the foregoing core and/or surface antigen epitopes is expressed as a fusion protein in the vaccinia viruses of the invention. "Fusion protein," as used herein, refers to a protein comprising an amino acid sequence from a first protein covalently linked via a peptide bond at its carboxy terminus to the amino terminus of an amino acid sequence from a second, different protein.

In one embodiment, a vaccine formulation of the invention contains a single type of recombinant vaccinia virus of the invention. In another embodiment, a vaccine formulation comprises a mixture of two or more recombinant viruses of the invention. Preferred viruses contain plural HBV antigens including but not limited to a combination of one of S, MS, LS, and full-length core, coreΔ8 (see Section 6.1.2 *infra*), preS1-coreΔ8, preS1-complete core, core-preS1*, core-preS2, or core-S* (the asterisks denoting the presence of an immunogenic fragment of the preceding antigen; hyphens denoting a fusion protein).

In various embodiments, the core and/or surface antigen epitopes can be expressed by the same recombinant virus, under the control of different promoters active in vaccinia virus. In a preferred aspect, two divergently oriented promoters are used. In a specific embodiment, a recombinant vaccinia virus expresses a surface-core (written in the amino to carboxy-terminal order) or core-surface fusion protein (*e.g.*, core-preS1, core-preS2, core-S, or preS1-core) under the control of a first promoter, and expresses a

surface antigen (*e.g.*, MS) under the control of a second promoter.

In another specific embodiment, a mixture of recombinant vaccinia viruses is provided, comprising a first virus capable of expressing core-preS1, a second virus capable of expressing core-preS2, and a third virus capable of expressing core-S. In another embodiment, such a mixture comprises a first virus capable of expressing core-preS1 or core-preS1*, a second virus capable of expressing core-preS2, and a third virus capable of expressing S. In another embodiment, such a mixture comprises a first virus capable of expressing core-preS1 or core-preS1*, and a second virus capable of expressing MS. In yet another embodiment, such a mixture comprises a first virus capable of expressing core-preS2, and a second virus capable of expressing S.

Other specific embodiments include the recombinant viruses disclosed in Section 6 *infra*.

The vaccine formulations of the present invention provide efficacious protection against HBV infection, because, as live viral vaccines, multiplication of the vaccine strain and expression of its HBV epitope(s) in the host leads to prolonged immune stimulus, and the combination of HBV immunogenic epitopes thereby expressed leads to protective immunity. The prolonged immune stimulus provided by the live viral vaccines of the invention is of similar kind and magnitude to that occurring in natural subclinical infections, and therefore, can confer substantial long-lasting immunity. The identity and structural context of the immunogenic epitopes expressed according to the invention provide protection against HBV infection. Thus, the present invention provides methods of prevention or treatment

of HBV infection and its clinical manifestations (hepatitis, hepatoma) comprising administering one or more of the recombinant vaccinia viruses of the invention.

Furthermore, the vaccine formulations of the invention provide one or more of the following benefits: stability for long periods without refrigeration; ease of production; low cost of production; ability to be administered by local workers without advanced medical training; and effectiveness in one dose. The vaccinia viruses which express an immunogenic fragment of more than one HBV protein cause immunological reactions against more components of HBV than can be presented by single subunit vaccines.

5.1. DNA SEQUENCES ENCODING HBV EPITOPES

Any DNA sequence which encodes an immunogenic (capable of provoking an immune response) HBV epitope, which produces protective immunity against hepatitis when expressed as a fusion or nonfusion protein in a recombinant vaccinia virus of the invention, alone or in combination with other HBV epitopes expressed by the same or a different vaccinia virus recombinant, can be isolated for use in the vaccine formulations of the present invention.

In a specific embodiment, epitopes of HBV antigens can be identified by virtue of their hydrophilicity, by carrying out a hydrophilicity analysis (Hopp and Woods, 1981, Proc. Natl. Acad. Sci. USA 78:3824) to generate a hydrophilicity profile. A hydrophilicity profile can be used to identify the hydrophobic and hydrophilic regions of a protein and the corresponding regions of the gene sequence which encode such proteins.

Preferred epitopes are those of the surface and core antigens described *supra*, specific embodiments of which are disclosed in examples section 6 herein (see also Table I). However, many different HBV surface and core antigen proteins or peptides can be expressed. These include but are not limited to naturally occurring membrane glycoproteins produced from recombinant vaccinia infected cells, which are glycosylated and may be transported to the cell surface. Peptides or proteins comprising surface antigen epitopes include but are not limited to an approximately 42,000 dalton glycoprotein consisting of preS1-preS2-S; an ~39,000 dalton unglycosylated protein consisting of preS1-preS2-S; an ~36,000 dalton glycoprotein (glycosylated at two sites) consisting of preS2-S; an ~33,000 dalton glycoprotein (glycosylated at one site) consisting of preS2-S; an ~27,000 dalton glycoprotein (glycosylated at one site) consisting of S; and an ~24,000 dalton unglycosylated protein consisting of S.

The immunogenic peptides and proteins expressed by the recombinant viruses of the invention can be secreted or not secreted by a host cell infected with the virus. Such proteins which are not secreted can be intracellular proteins or cell surface membrane proteins. In a preferred aspect, the HBV protein or peptide (containing the immunogenic epitope) expressed by the recombinant vaccinia virus is localized extracellularly (secreted) or is a cell surface protein. However, extracellular or cell surface localization is not required, since intracellular localization can also evoke an effective immune response (see, *e.g.*, Brown et al., 1987, J. Inf. Disc. 155:86).

The sequences encoding the immunogenic peptides or proteins are preferably present in single copies, but can also be present in multiple copies within the vaccinia virus genome. If multiple copies are present, care must be taken to ascertain that the recombinant virus stably maintains each of the multiple copies. In a preferred aspect, in an embodiment in which two copies are present, stability of the copies in the genome is maintained by localizing the individual copies relatively distant from each other in the genome with essential sequences between the copies, or, perhaps, by orienting the copies divergently. Stability of the copies in the genome can be confirmed by methods known in the art, *e.g.*, Southern analysis. In a specific embodiment, multiple copies are avoided by employing, in the recombinant vaccinia viruses of the invention, HBV sequences which do not overlap by (*i.e.*, do not contain an identical sequence of) more than 20 nucleotides, and preferably, do not overlap by more than 10 nucleotides. In a particular embodiment wherein a recombinant virus of the invention encodes a first and a second protein, each protein containing core and/or surface antigen epitopes, to maintain stability, the core and/or surface antigen sequences, or portions thereof displaying the antigenicity of such core and surface antigen sequences, respectively, in the first fusion protein are not present in the second fusion protein.

In a preferred aspect, the nucleotide sequence encoding the peptide or proteins comprising HBV immunogenic epitope(s) lacks introns, *e.g.*, is a cDNA sequence, since any introns will not be spliced out. Thus, for example, no introns containing stop codons in the desired reading frame may be contained in the HBV-encoding nucleotide sequence. The

nucleotide sequence preferably contains a ribosomal binding site, or is placed within an insertion vector (see Section 5.2) so as to be operatively linked to a ribosomal binding site.

The DNA sequence encoding the HBV epitope(s) can be obtained from any of numerous sources such as cloned HBV DNA, genomic HBV DNA, cDNA of HBV RNA, or chemically synthesized DNA (see, *e.g.*, U.K. Patent Publication No. GB 2034323A published June 4, 1980; Ganem and Varmus, 1987, Ann. Rev. Biochem. 56:651-693; Tiollais et al., 1985, Nature 317:489-495), and manipulated by recombinant DNA methodology well known in the art (see Sambrook et al., 1991, Molecular Cloning, A Laboratory Manual, 2d ed., Cold Spring Harbor Laboratory Press, New York).

Plasmid pAM6 contains the entire genome of HBV (subtype adw), and is available from the American Type Culture Collection (ATCC). In this plasmid, the HBV DNA has been inserted into the cloning vector pBR322 at a BamHI restriction enzyme site within the preS2 region to allow replication of the DNA in *E. coli*. Thus, the MS and LS antigen open reading frames (ORFs) in pAM6 are interrupted. In order to express the MS and LS antigens, it is necessary to reunite the surface antigen ORFs (*e.g.*, as described in the examples sections *infra*).

In order to generate HBV DNA fragments, the HBV DNA may be cleaved at specific sites using various restriction enzymes. Alternatively, one may use DNaseI in the presence of manganese, or mung bean nuclease (McCutchan et al., 1984, Science 225:626), to fragment the DNA, or the DNA can be physically sheared, as for example, by sonication. The linear DNA fragments can then be separated according to size by standard techniques, including, but not limited to,

agarose and polyacrylamide gel electrophoresis and column chromatography.

Any restriction enzyme or combination of restriction enzymes may be used to generate HBV DNA fragment(s) containing the desired epitope(s), provided the enzymes do not destroy the immunopotency of the encoded product. Consequently, many restriction enzyme combinations may be used to generate DNA fragments which, when inserted into an appropriate vector, are capable of directing the production of the peptide containing the HBV epitope(s).

Once the DNA fragments are generated, identification of the specific DNA fragment containing the desired HBV sequence may be accomplished in a number of ways. For example, if a small amount of the desired DNA sequence or a homologous sequence is previously available, it can be used as a labeled probe (*e.g.*, nick translated) to detect the DNA fragment containing the desired sequence, by nucleic acid hybridization. Alternatively, if the sequence of the derived gene or gene fragment is known, isolated fragments or portions thereof can be sequenced by methods known in the art, and identified by a comparison of the derived sequence to that of the known DNA or protein sequence. Alternatively, the desired fragment can be identified by techniques including but not limited to mRNA selection, making cDNA to the identified mRNA, chemically synthesizing the gene sequence (provided the sequence is known), or selection on the basis of expression of the encoded protein (*e.g.*, by antibody binding) after "shotgun cloning" of various DNA fragments into an expression system.

The sequences encoding HBV peptides to be expressed in recombinant vaccinia viruses according to the present invention, whether produced by recombinant DNA methods, chemical synthesis, or purification techniques, include but are not limited to sequences encoding all or part (fragments) of the amino acid sequences of HBV-specific antigens, as well as other derivatives and analogs thereof. In a specific embodiment, the derivative or analog is functionally active, *i.e.*, capable of exhibiting one or more functional activities associated with a full-length, wild-type HBV antigen (*e.g.*, core or surface antigen). In a preferred embodiment, such derivatives or analogs which have the desired immunogenicity or antigenicity (capable of being bound by an antibody to the HBV antigen) can be encoded. In another specific embodiment, a surface antigen derivative which retains the ability to assemble into 22 nm particles is encoded. In yet another specific embodiment, a core antigen derivative which retains the ability to assemble into core particles (see, *e.g.*, Cohen and Richmond, 1982, Nature 296:677-678) is encoded. In yet another embodiment, core and/or surface antigen derivatives which retain the ability to assemble into viral particles are encoded. Derivatives or analogs of HBV antigens can be tested for the desired activity by procedures known in the art, including but not limited to standard immunoassays.

In particular, HBV antigen derivatives can be made by altering the encoding HBV antigen nucleotide sequences by substitutions, additions or deletions that provide for functionally equivalent molecules. Due to the degeneracy of nucleotide coding sequences, other DNA sequences which encode substantially the same amino acid sequence as a native

HBV gene or portion thereof may be used in the practice of the present invention. These include but are not limited to nucleotide sequences comprising all or portions of HBV genes or cDNAs which are altered by the substitution of different codons that encode a functionally equivalent amino acid residue within the sequence, thus producing a silent change. For example, one or more amino acid residues within the sequence can be substituted by another amino acid of a similar polarity which acts as a functional equivalent, resulting in a silent alteration. Substitutes for an amino acid within the sequence may be selected from other members of the class to which the amino acid belongs. For example, the nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine. The polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine. The positively charged (basic) amino acids include arginine, lysine, and histidine. The negatively charged (acidic) amino acids include aspartic and glutamic acid.

Derivatives or analogs of HBV antigens include but are not limited to those peptides which are substantially homologous to an HBV antigen or fragments thereof, or whose encoding nucleic acid is capable of hybridizing to an HBV antigen-encoding nucleic acid sequence, *e.g.*, under conditions of high stringency (*e.g.*, 0.1 X SSC, 65°C).

The HBV antigen derivatives and analogs can be produced by various methods known in the art. For example, a cloned HBV gene sequence can be modified by any of numerous strategies known in the art (Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, 2d ed., Cold Spring Harbor Laboratory, Cold Spring

Harbor, New York). The sequence can be cleaved at appropriate sites with restriction endonuclease(s), followed by further enzymatic modification if desired, isolated, and ligated *in vitro*. In the production of the gene encoding a derivative or analog of an HBV antigen, care should be taken to ensure that the modified gene remains within the same translational reading frame as the antigen, uninterrupted by translational stop signals, in the gene region where the desired HBV epitope(s) are encoded.

Additionally, the HBV antigen-encoding nucleic acid sequence can be mutated *in vitro* or *in vivo*, to create and/or destroy translation, initiation, and/or termination sequences, or to create variations in coding regions and/or form new restriction endonuclease sites or destroy preexisting ones, to facilitate further *in vitro* modification. Any technique for mutagenesis known in the art can be used, including but not limited to, *in vitro* site-directed mutagenesis (Hutchinson, C., et al., 1978, J. Biol. Chem 253:6551), use of TAB® linkers (Pharmacia), etc.

In another specific embodiment, the encoded HBV antigen derivative is a chimeric, or fusion, protein comprising an HBV protein or fragment thereof fused to a non-HBV amino acid sequence. Such a chimeric protein is encoded by a nucleic acid encoding the HBV coding sequence joined in-frame to a non-HBV coding sequence. Such a chimeric product can be made by ligating the appropriate nucleic acid sequences encoding the desired amino acid sequences to each other by methods known in the art, in the proper coding frame.

A DNA sequence encoding an HBV epitope which is a hapten, *i.e.*, a molecule that is antigenic in

that it can react selectively with cognate antibodies, but not immunogenic in that it cannot elicit an immune response when administered without adjuvants or carrier proteins, can be isolated for use in the vaccine formulations of the present invention, since it is envisioned that, in particular embodiments, presentation by the vaccinia viruses of the invention can confer immunogenicity to the hapten expressed by the virus.

Once identified and isolated, the HBV DNA containing the sequence(s) of interest is then inserted into a recombinant vaccinia virus such that it can be expressed by such virus. Preferably, this is accomplished by first inserting the HBV DNA into a plasmid vector which is capable of subsequent transfer to a vaccinia virus genome by homologous recombination (see Section 5.2 *infra*).

5.2. CONSTRUCTION OF INSERTION VECTORS CONTAINING SEQUENCES WHICH ENCODE HBV SURFACE AND/OR CORE EPITOPE(S)

In a preferred aspect of the invention for constructing the recombinant vaccinia viruses, the desired DNA sequence encoding the HBV epitope(s) is inserted, using recombinant DNA methodology (see Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, 2d ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York) into an insertion (preferably, plasmid) vector flanked by (preferably) nonessential vaccinia DNA sequences, thus providing for subsequent transfer of its chimeric gene(s) into vaccinia virus by homologous recombination. The HBV sequences are placed in the vector such that they can be expressed under the control of a promoter functional in vaccinia virus.

Expression of foreign DNA in recombinant vaccinia viruses requires the positioning of promoters functional in vaccinia so as to direct the expression of the protein-coding HBV DNA sequences. Plasmid insertion vectors have been constructed to insert chimeric genes into vaccinia virus. One type of plasmid insertion vector is composed of: (a) a vaccinia virus promoter including the transcriptional initiation site; (b) one or more unique restriction endonuclease cloning sites located downstream from the transcriptional start site for insertion of foreign DNA fragments; (c) nonessential vaccinia virus DNA (such as the thymidine kinase (TK) gene) flanking the promoter and cloning sites which directs insertion of the chimeric gene into the homologous nonessential region of the virus genome; and (d) a bacterial origin of replication and antibiotic resistance marker for replication and selection in *E. coli*. Examples of such vectors are described by Mackett (Mackett et al., 1984. J. Virol. 49:857-864). The DNA encoding HBV epitope(s) is inserted into a suitable restriction endonuclease cloning site (b, above). Various derivatives of this prototype insertion vector can also be used, *e.g.*, with multiple restriction sites, with the β-galactosidase gene (*lacZ*) to allow selection of plaques by color, with inducible gene expression (*e.g.*, by inclusion of a bacterial lac repressor system), etc. (see, *e.g.*, Boyle et al., 1985, Gene 35:169-177; Chakrabarti et al., 1985, Mol. Cell. Biol. 5:3403-3409; Falkner et al., 1988, J. Virol. 62:1849-1854; Falkner et al., 1987, Nucl. Acids Res. 15:7192; Franke et al., 1985, Mol. Cell. Biol. 5:1918-1924; Panicali et al., 1986, Gene 47:193-199; Patel et al., 1988, Proc. Natl. Acad. Sci. USA 85:9431-9435; Fuerst et al., 1989, Proc. Natl. Acad.

Sci. USA 86:2549-2553). In addition to plasmid insertion vectors, insertion vectors based on single-stranded M13 bacteriophage DNA (Wilson et al., 1986, Gene 49:207-213) can be used.

The inserted HBV DNA should preferably not contain introns since intron sequences will not be removed during the viral life cycle, and insertion should preferably be so as to place the HBV coding sequences in close proximity to the promoter, with no other start codons in between the initiator ATG and the 5′ end of the transcript.

Preferably, a plasmid insertion vector such as described *supra*, with a convenient restriction site for insertion of the HBV DNA, is used. Alternatively, if the complementary restriction sites used to fragment the HBV DNA are not present in the cloning vector, the ends of the DNA molecules may be modified. Such modifications include producing blunt ends by digesting back single-stranded DNA termini or by filling the single-stranded termini so that the ends can be blunt-end ligated. Alternatively, any site desired may be produced by ligating nucleotide sequences (linkers) onto the DNA termini; these ligated linkers may comprise specific chemically synthesized oligonucleotides encoding restriction site recognition sequences. Other methods known in the art may be used.

The plasmid insertion vector should contain transcriptional and translational regulatory elements that are active in vaccinia virus. High expression levels in the derived recombinant virus can be obtained by using strong promoters or, less preferably, by inserting multiple copies of a single HBV sequence. In the use of multiple copies, care should be taken to confirm stable maintenance of the

copies in the vaccinia virus genome (see discussion *supra*).

As mentioned *supra*, the plasmid should be configured so that the HBV sequences are under the control of a promoter active in vaccinia virus. Such a promoter can be a native vaccinia promoter or a non-native promoter active in vaccinia virus. The promoter can be natural or synthetic. The plasmid can contain more than one promoter, each directing the expression of a different sequence encoding HBV epitope(s). In a preferred aspect, two such divergently oriented promoters are used. Promoters which can be used in the insertion vectors include but are not limited to the vaccinia virus thymidine kinase (TK) promoter, the 7.5K promoter (termed herein "p7.5"), a modified p7.5 (see *infra*), the 11K promoter (termed herein "p11"), the F promoter (Paoletti et al., 1984, Proc. Natl. Acad. Sci. USA 81:193-197), and various early and late vaccinia promoters (see Moss, 1990, *Virology*, 2d ed., ch. 74, Fields et al., eds., Raven Press, Ltd., New York, pp. 2079-2111). In the use of a native vaccinia virus promoter, a promoter active both early and late in the viral life cycle is preferred.

In a specific embodiment, the plasmid insertion vector contains (for eventual transfer into vaccinia virus) a T7 RNA polymerase coding sequence under the control of a promoter active in vaccinia virus, and the HBV coding sequences under the control of the T7 promoter. In another specific embodiment, a plasmid insertion vector contains a co-expression system consisting of divergently oriented promoters, one directing transcription of the HBV sequences, the other directing transcription of a reporter gene or selectable marker, to facilitate detection or

selection of the eventual recombinant vaccinia virus (see, *e.g.*, Fuerst et al., 1987, Mol. Cell. Biol. 5:1918-1924). In preferred aspects, the p7.5, p11, or modified versions thereof are employed, some of which are described in more detail below.

A strong vaccinia promoter, denoted here as "p7.5" (Cochran et al., 1985, J. Virol. 54:30-37), having the sequence provided in Figure 2 (SEQ ID NO:2), can be one of several vaccinia or vaccinia-like promoters used to direct expression of the HBV sequences. The p7.5 promoter is so named because it normally directs expression of a vaccinia polypeptide of 7.5 kilodaltons molecular weight. The transition from early to late events in viral replication is generally considered to be marked by the onset of viral DNA replication, and p7.5 contains elements which are active both early and late in viral replication, making it act as a constitutive promoter. Promoter p7.5 is contained in plasmid pGS53 (Fig. 3) (Chakrabarti et al., 1985, Mol. Cell. Biol. 5:3403-3409; kindly provided by Dr. Bernard Moss, National Institutes of Health). This plasmid contains fragments of the vaccinia thymidine kinase (TK) gene positioned so as to facilitate homologous recombination of the plasmid into that site in the vaccinia genome.

Another vaccinia promoter, p11, in nature directs expression of a vaccinia structural protein of 11 kilodaltons (European patent application 0198328, Hoffman LaRoche). p11, unlike the p7.5 promoter, is active only late in viral replication. The nucleotide sequence of p11 is shown in Figure 4 (SEQ ID NO:3). As described in Example 6, *infra*, a number of constructs were made using this promoter, both alone and in combination with p7.5, to direct expression of

HBV antigens. The 11 kD promoter is contained in a plasmid called pSC10 (Fig. 5, Chakrabarti et al., *supra;* kindly provided by Dr. Bernard Moss), which contains the gene coding for the bacterial enzyme β-galactosidase, *lacZ*, located downstream of the p11 promoter. In a specific embodiment, a p11-β-gal cassette is inserted into the middle of the vaccinia TK gene so that this expression unit can be inserted by *in vivo* recombination into the TK gene in the virus.

A strong vaccinia-like promoter used to direct expression of the hepatitis B antigens, termed herein "modified p7.5" has the nucleotide sequence set forth in Figure 6 (SEQ ID NO:4). The modified p7.5 promoter (constructed in the laboratory of Dr. Bernard Moss) is a strong synthetic promoter that is active both early and late in viral replication. The modified p7.5 promoter has a sequence based partly on the sequence of the p7.5 promoter. The modified p7.5 promoter is contained in plasmid pSC59 (Fig. 7; kindly provided by Dr. Bernard Moss).

Specific initiation signals are also required for efficient translation of inserted protein coding sequences. These signals include the ATG initiation codon and adjacent sequences. In cases where the entire HBV gene including its own initiation codon and adjacent sequences are inserted into the appropriate vectors, no additional translational control signals may be needed. However, in cases where only a portion of the gene sequence is inserted, exogenous translational control signals, including the ATG initiation codon, must be provided. The initiation codon must furthermore be in phase with the reading frame of the protein coding sequences to ensure translation of the entire insert. These

exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic.

In a preferred embodiment, the ATG initiation codon of the HBV antigen or portion thereof is "upregulated," *i.e.*, contains a sequence upstream of that encoding the initiator methionine that matches a consensus sequence associated with strong ribosome binding. In a specific embodiment, the ATG codon of MS is upregulated.

As described *supra*, the plasmid insertion vector contains at least one set of HBV coding sequences operatively linked to a promoter, flanked by sequences preferably nonessential for vaccinia viral replication. Such nonessential sequences include but are not limited to the TK gene (Mackett et al., 1984, J. Virol. 49:857-864), the vaccinia HindIII-F DNA fragment (Paoletti et al., 1984, Proc. Natl. Acad. Sci. USA 81:193-197), the vaccinia growth factor gene situated within both terminal repeats (Buller et al., 1988, J. Virol. 62:866-874), the N2 and M1 genes (Tamin et al., 1988, Virology 165:141-150), the M1 subunit of the ribonucleotide reductase gene in the vaccinia HindIII-I DNA fragment (Child et al., 1990, Virology 174:625-629), the vaccinia hemagglutinin (Shida et al., 1988, J. Virol. 62:4474-4480), vaccinia 14 kD fusion protein gene (Rodriguez et al., 1989, Proc. Natl. Acad. Sci. USA 86:1287-1291), etc. (see also Buller and Palumbo, 1991, Microbiol. Rev. 55(1):80-122). TK sequences are preferred for use; use of such sequences results in the generation of TK^- recombinant viruses. TK^- recombinants have been shown to be attenuated relative to the parent strain (Buller et al., 1985, Vaccines 85:163), which may be an

advantage from a safety standpoint in the development of new vaccinia-based vaccines.

In an alternative embodiment, because vaccinia recombinants containing HBV sequences within coding regions such as TK may be overly attenuated, the HBV-encoding sequences are inserted into a region of the vaccinia genome that does not encode protein, *e.g.*, an intergenic region. In addition, insertion into such a non-coding region in conjunction with insertion into a nonessential coding region such as TK might enable the generation of recombinant vaccinia viruses with multiple insertions which are no more attenuated than those viruses with a single insertion only in a nonessential coding region. The strategy for choosing a non-coding region is based on several parameters, one of which is the length of the region, *i.e.*, the distance between surrounding genes. A larger distance is preferable to a smaller distance in order to avoid regulatory regions necessary for the expression of surrounding genes. Another parameter is the function of the surrounding genes, nonessential genes being preferable to essential genes. A sequence and map of the wild-type vaccinia Copenhagen strain is found in Goebel et al., 1990, Virology 179:247-266. For strains which have not been sequenced, the sequence of the chosen non-coding region is obtained or determined according to conventional nucleotide sequence techniques.

By way of example but not limitation, construction of an insertion vector with intergenic flanking regions can be carried out as follows:

After an intergenic region is selected as a candidate nonessential non-coding region for insertion of foreign DNA, a pair of polymerase chain reaction (PCR) primers having nucleotide sequences

corresponding to a sequence within the chosen intergenic region is synthesized or obtained commercially. Primers are preferably chosen such that there is between 100 and 300 nucleotides between primers of a pair. 300 nucleotides has been designated herein as an upper length limit because the largest known non-coding region in the Copenhagen strain is approximately 400 nucleotides in length. In order to avoid potential regions of control of surrounding genes, it is preferable to leave a minimum of 50 nucleotides untouched on either side of the insertion region. Longer DNA sequences can also be amplified if primers hybridizing within the surrounding genes are used. Primers are designed so as to hybridize to complementary strands and allow amplification of the region between them. To facilitate later cloning, it is desirable to build a convenient restriction site into the primer sequence. Examples of sets of primers which may be used for the following non-coding regions, along with convenient restriction sites, are as follows.

TABLE 1

Non-coding Region	Primers	SEQ ID NO:	SITE
F14L-F15L	5'GTTGGTAGAATTCCAATTAT3' 5'AAGAATGATCGATACAGTTT3'	50 51	EcoRI ClaI
C12L-C11R	5'GTACCCCGAATTCATACTTA3' 5'AATAATATATCGATAATTGT3'	52 53	EcoRI ClaI
A53R-A55R	5'GTTGGAATTCCGCTACTGAT3' 5'TAACCAAGTATCGATATAAT3'	54 55	EcoRI ClaI

Viral DNA is isolated from the wild type, *e.g.*, Wyeth virus according to known procedures. The primers are then hybridized to the viral DNA under non-denaturing conditions. PCR (see Section 5.3) is then performed on the isolated DNA using the pair of primers corresponding to the chosen intergenic region. The amplified DNA is then digested with restriction enzymes which cut within the primer sequences and the resulting fragment is gel purified. The gel purified DNA is cloned into an appropriate vector using standard cloning techniques.

The cloned non-coding region is digested with a restriction enzyme at a site within the non-coding region. Noncoding regions which can potentially be used, with their approximate size and a convenient restriction site for insertion, are set forth in Table 2.

TABLE 2

Non-coding Region	Approximate Band Size of Non-coding Region	Internal Restriction Site
F14L-F15L	210 bp	SnaBI
C12L-C11R	260 bp	BclI
A53R-A55R	240 bp	SnaBI

In order to clone the desired HBV sequences downstream of vaccinia promotors in the vector, the desired promoter/HBV DNA construct is inserted within the cloned non-coding region using the appropriate restriction enzymes and sites. Preferably, the HBV sequences are inserted so as to be flanked by at least

about 100 bp of the noncoding vaccinia DNA. Most preferably the sequences are inserted into the vector so as to be surrounded by about 500 bp of vaccinia DNA on each side, which may include some DNA from nonessential vaccinia genes.

5.3. CONSTRUCTION AND IDENTIFICATION OF RECOMBINANT VACCINIA VIRUSES ENCODING HBV EPITOPE(S)

Recombinant vaccinia viruses are preferably produced by transfection of the recombinant insertion vectors containing the HBV sequences into cells previously infected with vaccinia virus. Alternatively, transfection can take place prior to infection with vaccinia virus. Homologous recombination takes place within the infected cells and results in the insertion of the foreign gene into the viral genome, in the region corresponding to the insertion vector flanking regions. The infected cells can be screened using a variety of procedures such as immunological techniques, DNA plaque hybridization, or genetic selection for recombinant viruses which subsequently can be isolated. These vaccinia recombinants preferably retain their essential functions and infectivity and can be constructed to accommodate up to approximately 35 kilobases of foreign DNA.

Transfections may be performed by procedures known in the art, for example, a calcium chloride-mediated procedure (Mackett et al., 1985, The construction and characterization of vaccinia virus recombinants expressing foreign genes, in *DNA Cloning*, Vol. II, Rickwood and Hames (eds.), IRL Press, Oxford-Washington, D.C.) or a liposome-mediated procedure (Rose et al., 1991, Biotechniques 10:520-525).

Where, as is preferred, flanking TK sequences are used to promote homologous recombination, the resulting recombinant viruses thus have a disrupted TK region, permitting them to grow on a TK^- host cell line such as Rat2 (ATCC Accession No. CRL 1764) in the presence of 5-bromo-2'-deoxyuridine (BUDR), under which conditions non-recombinant (TK^+) viruses will not grow. TK^- recombinants have been shown to be attenuated relative to the parent strain (Buller et al., 1985, Infectious vaccinia virus TK^- recombinants that express foreign genes are less virulent than wild-type virus in mice, in *Vaccines 85, Molecular and Chemical Basis of Resistance to Parasitic, Bacterial and Viral Disease*, R.A. Lerner, Chanock, R.M., and Brown, F. (eds.), Cold Spring Harbor Laboratory, New York, pp. 163-167), which may be an advantage from a safety standpoint in the development of new vaccinia-based vaccines.

In another embodiment, recombinant vaccinia viruses of the invention are made by *in vitro* cloning, and then packaging with a poxvirus sensitive to a selection condition, rather than by homologous recombination. For example, the HBV DNA sequences can be inserted into vaccinia genomic DNA using standard recombinant DNA techniques *in vitro*; this recombinant DNA can then be packaged in the presence of a "helper" poxvirus such as a temperature sensitive vaccinia virus mutant or a fowlpox virus which can be selected against under the appropriate conditions.

Various vaccinia virus strains known in the art can be used to generate the recombinant viruses of the invention. Characterized strains of low virulence are preferred. A preferred vaccinia virus is the New York City Department of Health Laboratories strain, prepared by Wyeth (available from the American Type

Culture Collection (ATCC), Accession No. VR-325). Other vaccinia strains include but are not limited to the Elstree and Moscow strains, the strain of Rivers (CV-1 and CV-2), and the LC16m8 strain of Hashizume.

Selection of the recombinant vaccinia virus can be by any method known in the art, including hybridization techniques (*e.g.*, using HBV DNA sequences as a hybridization probe), immunological techniques (*e.g.*, assay for binding to antibodies recognizing the encoded HBV epitope(s)), etc. In a preferred aspect where TK flanking sequences are used in the insertion vector, selection is for TK^- recombinants, as described above; screening for the correct recombinant is then carried out by molecular analyses as described *infra*. In many preferred aspects, the method of choice for selection is dictated by the choice of insertion vector used to generate the recombinant viruses. For example, if a coexpression system vector is used, containing a reporter gene or selectable marker, detection of recombinants is carried out by screening for the reporter gene expression or selecting for selectable marker expression. In an embodiment where the *lacZ* gene is incorporated, color screening methods known in the order can be used to detect expression of β-galactosidase. Incorporation of neomycin or gpt genes allows selection by antibiotic resistance. Plaque size can also be used for screening. Many other methods will be known to the skilled artisan and can be used.

The selected recombinant vaccinia virus is then generally plaque-purified (preferably by at least three rounds of purification from a single viral plaque), and subjected to molecular analyses to verify its identity and purity. Preferably, two types of

molecular analyses are carried out: (i) nucleic acid analyses; and (ii) assays for expression of the encoded HBV epitope(s).

For nucleic acid analyses, to verify that the recombinant vaccinia virus contains the DNA sequences encoding HBV epitope(s), hybridization assays are preferably carried out using a labelled nucleic acid containing HBV sequences as a hybridization probe. For example, plaque-lift hybridizations (see Mackett et al., 1985, in *DNA Cloning*, Vol. II, Rickwood and Hames (eds.), IRL Press, Oxford-Washington, D.C.), Southern hybridizations (Southern, 1975, J. Mol. Biol. 98:503-517; see Section 8.1. *infra*), and Northern hybridizations (Freeman et al., 1983, Proc. Natl. Acad. Sci. USA 80:4094-4098) can be carried out. Other methods of nucleic acid analysis which can be employed are restriction endonuclease mapping (Sambrook et al., 1991, Molecular Cloning, A Laboratory Manual, 2d ed., Cold Spring Harbor Laboratory Press, New York), and DNA sequence analysis (*e.g.*, according to Maxam and Gilbert, 1980, Meth. Enzymol. 65:499-560; Sanger et al., 1977, Proc. Natl. Acad. Sci. USA 74:5463; or Tabor and Richardson, U.S. Patent No. 4,795,699; or by use of an automated DNA sequenator). If desired, polymerase chain reaction (PCR; U.S. Patent Nos. 4,683,202, 4,683,195 and 4,889,818; Gyllenstein et al., 1988, Proc. Natl. Acad. Sci. USA 85:7652-7656; Ochman et al., 1988, Genetics 120:621-623; Loh et al., 1989, Science 243:217-220) can be done prior to sequencing in order to amplify the region of interest.

Southern blot hybridization can also be carried out using a hybridization probe containing the desired vaccinia sequences to ensure that the HBV

sequences have been inserted into the desired region of the vaccinia genome.

Protein analyses, to confirm expression of the encoded HBV epitope(s) by the selected recombinant vaccinia virus, can be carried out by any method known in the art, including functional or immunological assays, and are preferably accomplished by immunoassay methods employing antibodies to the encoded HBV epitope(s). Various immunoassays known in the art can be used, including but not limited to competitive and non-competitive assay systems using techniques such as radioimmunoassays, ELISA (enzyme linked immunosorbent assay), "sandwich" immunoassays, immunoradiometric assays, gel diffusion precipitin reactions, immunodiffusion assays, *in situ* immunoassays (using colloidal gold, enzyme or radioisotope labels, for example), western blots, immunoprecipitation reactions, agglutination assays (*e.g.*, gel agglutination assays, hemagglutination assays), complement fixation assays, immunofluorescence assays, protein A assays, and immunoelectrophoresis assays, etc. In one embodiment, antibody binding is detected by detecting a label on the primary antibody. In another embodiment, the primary antibody is detected by detecting binding of a secondary antibody or reagent to the primary antibody. In a further embodiment, the secondary antibody is labelled. Many means are known in the art for detecting binding in an immunoassay and are envisioned for use.

In specific embodiments, western blot analyses can be carried out as described in Section 8.2. *infra*. ELISA can be carried out as described in Section 8.3 *infra*. Several immunoassays for HBV S or e antigens are commercially available (Abbott Laboratories, No. Chicago, IL).

Recombinant viruses containing a combination of epitopes from core and surface elements are a preferred embodiment of the invention and can be generated using the plasmids described in the examples sections *infra*, and homologous recombination.

5.4. DETERMINATION OF VACCINE EFFICACY

Immunopotency of the HBV epitope(s) in their live vaccinia vaccine formulation can be determined by monitoring the immune response of test animals following immunization with the recombinant vaccinia virus(es) expressing the HBV epitope(s). Generation of a humoral response may be taken as an indication of a generalized immune response, other components of which, particularly cell-mediated immunity, may be important for protection against HBV. Test animals may include mice, rabbits, chimpanzees and eventually human subjects. HBV can be made to infect chimpanzees experimentally, although they do not develop chronic infections. However, since chimpanzees are a protected species, the antibody response to an HBV vaccine of the invention can first be studied in a number of smaller, less expensive animals, with the goal of finding one or two best candidate viruses or best combinations of viruses to use in chimpanzee efficacy studies.

Methods of introduction of the vaccine may include oral, intracerebral, intradermal, intramuscular, intraperitoneal, intravenous, subcutaneous, intranasal or any other standard routes of immunization. The immune response of the test subjects can be analyzed by various approaches such as: the reactivity of the resultant immune serum to HBV antigens, as assayed by known techniques, *e.g.*, enzyme linked immunosorbent assay (ELISA),

immunoblots, radioimmunoprecipitations, etc.; or protection from HBV infection and/or attenuation of hepatitis symptoms in immunized hosts.

In a specific embodiment, recombinant vaccinia virus vaccines of the invention can be prepared in large quantities for animal studies. By way of example, this is done by infecting roller bottles containing 10^8 to 10^9 cells with virus, and harvesting after several days of infection. The cells are pelleted, washed, and then disrupted by resuspension and homogenization in a hypotonic buffer. Nuclei and cellular debris are then pelleted and the virus-containing supernatant is sonicated to break up any virus aggregates. The viruses are then pelleted through a cushion of 36% sucrose in a hard spin, followed by banding in a 24-40% sucrose gradient. The purified virus is then harvested by a hard spin, resuspended in a small volume, and titrated to determine viral concentration and yield. This process may be scaled up using a tangential flow filtration apparatus, and a continuous flow homogenizer.

As one example of suitable animal testing, HBV vaccines of the invention may be tested in rabbits for the ability to induce an antibody response to HBV antigens. Male specific-pathogen-free (SPF) young adult New Zealand White rabbits may be used. The test group of rabbits each receives approximately 5×10^8 pfu (plaque forming units) of the vaccine. A control group of rabbits receives an injection in 1 mM Tris-HCl pH 9.0 of the parental non-recombinant vaccinia virus or of a recombinant virus which does not express HBV antigens.

Blood samples may be drawn from the rabbits every one or two weeks, and serum analyzed for antibodies to the HBV antigens using, *e.g.*, a

radioimmunoassay (Abbott Laboratories). The presence of antibodies specific for core protein or the surface antigens may be assayed using an ELISA. The sera may also be analyzed for antibodies to vaccinia, *e.g.*, in an enzyme-linked immunoassay. Because rabbits may give a variable response due to their outbred nature, it may also be useful to test the vaccines in mice.

Mice respond differently to HBV antigens depending on their H-2 (histocompatibility) type (Milich and Chisari, 1982, J. Immunol. 129:320-325). Because there is little information available regarding the expression of HBV antigens from recombinant vaccinia viruses in mice, an HBV-responsive strain of mouse must be chosen, and an appropriate dose of vaccinia virus must be administered to render the mouse strain responsive to recombinant viruses of the invention (see Section 8.4, *infra*).

Patas monkeys may be used to test for immunogenicity of HBV vaccine formulation, although challenge experiments cannot be carried out since they are resistant to HBV infection. Chimpanzees may be tested for HBV vaccine efficacy (*e.g.*, challenge experiments), since they can be infected by the virus. In a specific embodiment, monkeys each receive intradermally approximately 5×10^8 pfu of recombinant Wyeth virus. A control monkey receives Wyeth (control) virus intradermally. Blood is drawn weekly for 12 weeks, and serum is analyzed for antibodies to vaccinia, core, and S.

5.5. VACCINE FORMULATION AND ADMINISTRATION

The purpose of this embodiment of the invention is to formulate a vaccine in which the immunogen is one or several recombinant vaccinia

virus(es) that express(es) HBV epitope(s) so as to elicit a protective immune (humoral and/or cell mediated) response against HBV infections for the prevention of hepatitis, hepatoma and/or other undesirable correlates of HBV infection.

Many methods may be used to introduce the vaccine formulations of the invention; these include but are not limited to oral, intradermal, intramuscular, intraperitoneal, intravenous, subcutaneous, intranasal routes, and via scarification (scratching through the top layers of skin, *e.g.*, using a bifurcated needle). In a preferred embodiment, scarification is employed; intravenous administration is disfavored, since it is preferred to have the virus replicate in the skin and not spread systemically.

The vaccine formulations of the invention comprise the recombinant vaccinia virus and a pharmaceutically acceptable carrier or excipient. The recombinant virus can be replicating or nonreplicating, although replicating is preferred. Similarly, an enveloped virus is preferred over the non-enveloped form. Pharmaceutically acceptable carriers include but are not limited to saline, buffered saline, dextrose, water, glycerol, sterile isotonic aqueous buffer, and combinations thereof. The formulation should suit the mode of administration.

The composition, if desired, can also contain minor amounts of wetting or emulsifying agents, or pH buffering agents. The composition can be a liquid solution, suspension, emulsion, tablet, pill, capsule, sustained release formulation, or powder. Oral formulation can include standard carriers such as pharmaceutical grades of mannitol,

lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate, etc.

Generally, the ingredients are supplied either separately or mixed together in unit dosage form, for example, as a dry lyophilized powder or water free concentrate in a hermetically sealed container such as an ampoule or sachette indicating the quantity of active agent. Where the composition is administered by injection, an ampoule of sterile diluent can be provided so that the ingredients may be mixed prior to administration.

In a specific embodiment, the lyophilized recombinant vaccinia virus of the invention is provided in a first container; a second container comprises diluent consisting of an aqueous solution of 50% glycerin, 0.25% phenol, and an antiseptic (*e.g.*, 0.005% brilliant green).

The precise dose of virus to be employed in the formulation will also depend on the route of administration, and the nature of the patient, and should be decided according to the judgment of the practitioner and each patient's circumstances according to standard clinical techniques. However, a suitable dosage range for administration via scarification with a bifurcated needle is 10^6 to 10^9 pfu/ml (see, *e.g.*, Cooney et al., 1991, The Lancet 337:567-572, regarding recombinant vaccinia viruses expressing HIV proteins). Boosting is possible but not preferred. If boosting is desired, it may be preferable to boost with the HBV antigen in purified form rather than using a recombinant virus of the invention, since long-lasting immunity to vaccinia virus may limit the utility of recombinant vaccinia virus for repetitive boosting (see *id.*). Effective

doses may also be extrapolated from dose-response curves derived from animal model test systems.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the vaccine formulations of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration.

5.6. USE OF ANTI-HBV ANTIBODIES GENERATED BY THE VACCINES OF THE INVENTION

The antibodies generated against HBV by immunization with the recombinant viruses of the present invention also have potential uses in diagnostic immunoassays, passive immunotherapy, and generation of antiidiotypic antibodies.

The generated antibodies may be isolated by standard techniques known in the art (*e.g.*, immunoaffinity chromatography, centrifugation, precipitation, etc.) and used in diagnostic immunoassays. The antibodies may also be used to monitor treatment and/or disease progression. Any immunoassay system known in the art, such as those listed *supra*, may be used for this purpose including but not limited to competitive and noncompetitive assay systems using techniques such as radioimmunoassays, ELISA (enzyme-linked immunosorbent assays), "sandwich" immunoassays, precipitin reactions, gel diffusion precipitin reactions, immunodiffusion assays, agglutination assays, complement-fixation assays, immunoradiometric assays,

fluorescent immunoassays, protein A immunoassays and immunoelectrophoresis assays, to name but a few.

The vaccine formulations of the present invention can also be used to produce antibodies for use in passive immunotherapy, in which short-term protection of a host is achieved by the administration of pre-formed antibody directed against a heterologous organism (in this case, HBV).

The antibodies generated by the vaccine formulations of the present invention can also be used in the production of antiidiotypic antibody. The antiidiotypic antibody can then in turn be used for immunization, in order to produce a subpopulation of antibodies that bind the initial antigen of the pathogenic microorganism (Jerne, 1974, Ann. Immunol. (Paris) 125c:373; Jerne, et al., 1982, EMBO J. 1:234).

5.7. THERAPEUTIC USE OF THE RECOMBINANT VACCINIA VIRUSES OF THE INVENTION

In this embodiment, the recombinant vaccinia viruses of the invention, encoding HBV immunogenic epitope(s), are administered therapeutically, for the treatment of hepatitis or other undesirable correlates of HBV infection. Administration of such viruses, *e.g.*, to neonates and other human subjects, can be used as a method of immunostimulation, to boost the host's immune system, enhancing cell-mediated and/or humoral immunity, and facilitating the clearance of HBV. The viruses of the invention can be administered alone or in combination with other anti-viral therapies, including but not limited to α-interferon and vidarabine phosphate.

6. CONSTRUCTION OF VECTORS ENCODING HBV EPITOPES

Some of the vectors which were constructed to contain DNA sequences encoding HBV surface antigen and/or core epitopes, under the control of a promoter functional in vaccinia virus, are listed in Table 3, along with their encoded antigens and the promoters operably linked to the HBV DNA. These vectors and their construction are described in the subsections below.

TABLE 3

Plasmids Containing Sequences Encoding HBV Epitopes Under the Control of a Promoter Functional in Vaccinia Virus

PLASMID	PROMOTER		
	p7.5	p11	modified p7.5
pSC10		β-galactosidase	
pGS53/S	S		
pGS53/S2	MS		
pGS53/S1	LS		
pHTL-5		core(full-length)	
pHTL-6		core(full-length)	S
pHTL-7		core(full-length)	LS
pHTL-8			S
pHTL-9			LS
pHTL-10			MS
pHTL-10M			MS (upreg. ATG)
pHTL-11		core(full-length)	MS
pHTL-12		core(Δ8)	MS
pHTL-13	core(full-length)	S	

PLASMID	PROMOTER		
	p7.5	p11	modified p7.5
pHTL-14	preS1-core(Δ8)		MS
pHTL-15	core(full-length)		S
pHTL-17	core(full-length)		MS
pHTL-18	preS1-core(full-length)		MS
pHTL-23		preS1-core(Δ8)	
pHTL-24		preS1-core(full-length)	
pHTL-25			core(full-length)
pHTL-26			core-preS1*
pHTL-27			core-preS2
pHTL-28			core-S*
pHTL-30	core(Δ8)		MS (upreg. ATG)
pHTL-31	preS1-core(Δ8)		MS (upreg. ATG)
pHTL-32	preS1-core(Δ8)		
pHTL-33	core(full-length)		MS (upreg. ATG)
pHTL-34	core(full-length)		S
pHTL-35	core(full-length)		LS
pHLT-36	preS1-core(full-length)		MS (upreg. ATG)
pRO-10		S	
pRO-11	MS	S	
pRO-13B	LS		
pRO-16	core(Δ8)	S	
pRO-16M	core(Δ8)		
pRO-17	MS	core(Δ8)	
pRO-18	LS	core(Δ8)	
pRO-19		core(Δ8)	

PLASMID	PROMOTER		
	p7.5	p11	modified p7.5
pRO-19/24		core(full-length)	
pRO-21	MS	preS1-core(Δ8)	
pRO-22	MS	core(full-length)	
pRO-23	MS	preS1-core(full-length)	

Other vectors which were constructed, and which have utility in generating the recombinant viruses of the invention upon further manipulation (with respect to the cloning vectors), or with a plasmid that provides T7 polymerase (with respect to HBV sequences operably linked to a T7 promoter) are also described *infra*.

6.1. SUBCLONING OF HBV GENES

The HBV genome, showing the core and surface antigen-encoding regions, is illustrated schematically in Figure 8. Using standard recombinant DNA techniques (see, for example, Sambrook et al., 1989, *Molecular Cloning, A Laboratory Manual*, 2nd ed., Cold Spring Harbor Press, New York), each of the four HBV genes were subcloned separately from the cloned HBV genome and placed in individual plasmids such that multiple restriction enzyme digestion sites surrounded each gene and provided for easy manipulation. A method used to make these subclones is described below. Unless otherwise mentioned, all DNA-modifying enzymes were obtained commercially from New England Biolabs (Beverly, MA) or BRL Life Technologies, Inc. (Gaithersburg, MD), and were used according to the manufacturer's instructions.

6.1.1. SURFACE ANTIGENS

The HBV genome (subtype adw) is contained in plasmid pAM6 (Moriarty et al., 1981, Proc. Natl. Acad. Sci. USA 78:2606-2610, available from the American Type Culture Collection (ATCC) Accession No. 45020, Fig. 9), a pBR322-based vector that is replicable in *E. coli*. Since the MS/LS open reading frames (ORFs) in pAM6 are interrupted by pBR322 DNA, these ORFs were resected using the 2883 bp cloning vector pT7T318 (Pharmacia cat. no. 27-3512; Fig. 10), which contains multiple restriction sites.

The 5' end of the open reading frame coding for the three surface antigens, which is bounded by BamHI and BglII restriction endonuclease sites in pAM6, a fragment of length 799 bp, was removed from pAM6 and inserted into BamHI-linearized pT7T318 to generate plasmid pRO-01 (Fig. 11). Briefly, pAM6 plasmid DNA was double digested with BamHI and BglII enzymes, and the fragments generated by the digestion were electrophoretically separated on an 0.8% agarose gel with appropriate size markers in 1X TBE (89.2 mM Tris pH 8.0, 89 mM boric acid, 2 mM EDTA), and the fragment corresponding to the 799 bp insert containing preS1 and part of preS2 was excised from the ethidium-bromide stained gel. This fragment was eluted from the gel by electrophoresis into a dialysis bag in 0.5X TBE and ligated, using T4 DNA ligase, to the pT7T318 DNA which had been linearized by BamHI digestion, similarly gel-purified, and treated with bacterial alkaline phosphatase (BAP) to prevent recircularization of the vector. A fraction of the ligation mix was transformed into *E. coli* (strain DH5α $F'I^Q$, available from BRL Life Technologies, Inc., cat. no. 8288SA) which had previously been made competent to take up the DNA by standard techniques

(Sambrook et al., *supra*). The transformed cells were plated on plates of LB agar plus 50 μg/ml ampicillin to select for cells which had been transformed with the plasmid. Small 1 ml cultures of LB broth (10 g/l bactotryptone, 5 g/l bacto-yeast extract, 10 g/l NaCl pH 7.5) plus 50 μg/ml ampicillin were then grown. DNA was prepared from these cultures by a boiling lysis miniprep protocol (Sambrook et al., *supra*) followed by digestion and electrophoresis on an agarose gel to screen for a colony containing an insert of the expected size and orientation. Specifically, the size and desired orientation of the insert was verified by digestion with PstI, which cuts in the polylinker of pT7T318 and within the insert, 5 bp from the 3' end of the preS2 region. In the desired orientation, digestion of pRO-01 with PstI yields fragments of 812 bp and 2870 bp, consistent with the PstI site of the pT7T318 polylinker being at the N-terminus of the open reading frame. Large quantities of the desired DNA and a glycerol stock of the plasmid-harboring *E. coli* were prepared by standard methods from a 250 ml LB broth plus ampicillin culture of the chosen colony.

The 3' end of the ORF coding for the three surface antigens was manipulated using similar techniques to connect it to the 5' end of the ORF. Specifically, the 3'-end fragment with length 1372 bp wa excised from pAM6 by BamHI digestion and purified from an agarose gel. This fragment was inserted into BamHI-linearized and bovine alkaline phosphatase (BAP)-dephosphorylated pRO-01, a fragment of 3683 bp, to generate pRO-02 (Fig. 12), which contains an intact open reading frame capable of encoding LS, MS and S antigens. The correct orientation of the insert in pRO-02 was verified by double digestion with PstI and

XbaI to yield a restriction pattern consistent with fragments of sizes 4018, 800, 224 and 12 bp.

In order to achieve a high level of expression of the genes encoding the LS, MS, and S antigens, each gene was engineered so as to be under the transcriptional control of a vaccinia or vaccinia-like promoter, *i.e.*, a promoter which is recognized by vaccinia virus RNA polymerase. In nature, the 3' end of the vaccinia promoter is often very close to, in the case of early promoters, or superimposed on, in the case of late promoters, the translation initiation codon of the ORF (1990, *Virology*, Fields, B.N., ed., Raven Press, New York), so we deemed this proximity also desirable in recombinant constructs. With this goal in mind, the individual S, MS, or LS ORFs were subcloned back into pT7T318 to place restriction endonuclease sites close to the initiating ATG of each ORF. Using these sites, the vaccinia or vaccinia-like promoters were placed very close to the initiating ATG of each ORF.

The translational initiation codons of the MS and S ORFs are contained within a 321 bp HincII fragment of pRO-02. This HincII fragment was isolated from pRO-02 and cloned into HincII-linearized and BAP-treated pT7T318 to generate pRO-03 (Fig. 13). The correct orientation of the inserted fragment was verified by digestion of pRO-03 with BamHI to yield a restriction pattern consistent with fragments of 201 bp and 3003 bp, indicating that the insert had been made such that the BamHI site in the pT7T318 polylinker was located at the 3' end of the insert.

From pRO-03, the initiating ATG of the S ORF was subcloned so as to be in close proximity to useful restriction sites. Specifically, pRO-03 was digested with NlaIII, which digests the plasmid in many places

(11 sites) including within the MS-encoding region one base pair upstream of the initiating ATG of the S ORF. NlaIII-digested pRO-03 was also digested with HincII, which cuts the plasmid at the 3' end of the S-fragment insert. The 67 bp ATG-containing fragment generated by this NlaIII/HincII digestion was then isolated from a 5.5% polyacrylamide gel run with appropriate size standards and then cloned into the vector backbone of pT7T318 which had been double digested with SphI and HincII, generating pRO-05 (Fig. 14), which contains the initiating ATG of the S ORF. The presence of the insert in pRO-05 was verified by double digestion with HindIII and EcoRI to yield a restriction pattern consistent with fragments of lengths 103 and 2871 bp. The orientation of the insert in pRO-05 was verified by restriction digestion to be such that the HincII site at the carboxyl-terminus of the inserted ORF is retained.

In a similar manner, the initiating ATG of MS was subcloned so as to be closely surrounded by restriction sites. Specifically, pRO-03 was digested with HaeIII, which cuts pRO-03 at 12 sites including one just two base pairs 5' of the initiating ATG of MS, together with HincII. The resulting 232 bp fragment containing the 5' end of the MS ORF was isolated and subcloned into the 2866 bp vector backbone of pT7T318 which had been double digested by SmaI and HincII, BAP-treated, and gel-purified, generating pRO-06 (Fig. 15). The size and orientation of the insert were verified by double digestion with HincII and EcoRI to yield a restriction pattern consistent with fragments of length 248 bp and 2850 bp.

The 3' end of the S antigen coding region was then added to the S ORF in pRO-05 by HincII and

StuI double digestion of pRO-02 and isolation of the 750 bp fragment coding for this C-terminal region. This fragment was inserted into the 2918 bp HincII and SmaI double digested vector fragment of pRO-05 to generate pT7T3/S (Fig. 16). Restriction analysis was used to confirm that the orientation of the insert in pT7T3/S was such that the HincII site within the S ORF had been recreated. pT7T3/S thus contains DNA encoding the entire open reading frame of the S antigen, beginning with the initiating ATG.

In a similar manner, the 3' end of the S antigen was added to the MS ORF in pRO-06 by double digesting pRO-02 with HincII and SphI, isolating the fragment corresponding to the 1016 bp 3' end, and inserting this fragment into the 3085 bp vector fragment of HincII and SphI double digested pRO-06, generating pT7T3/S2 (Fig. 17). Restriction analysis was used to verify that the insert's orientation in pT7T3/S2 was such that the HincII site within the S region had been recreated. pT7T3/S2 created in this manner contains DNA encoding the entire open reading frame of the MS antigen, beginning with the initiating ATG of MS.

A plasmid construct containing the LS open reading frame closely surrounded by restriction endonuclease sites was constructed using a similar procedure. The 509 bp HincII fragment of pAM6 containing the LS ATG was subcloned into HincII-linearized pT7T318, generating pLEH-01 (Fig. 18). pLEH-01 was then double digested with NlaIII, which cuts 10 times including at a site 1 base pair 5' of the initiating ATG of LS, and with HincII at a site approximately 230 base pairs 3' of the ATG. This fragment of 234 bp, containing the 5' end of the LS ORF, was subcloned into the 2869 bp vector backbone of

pT7T318 which had been double digested with SphI and HincII to give pLEH-02 (Fig. 19). Restriction analysis was used to confirm the size of the insert in pLEH-02 and that the insert was oriented so as to recreate the HincII site. The 132 bp middle fragment of the LS ORF stretching from within preS1 to within preS2 was then subcloned from HincII/BamHI-cut pRO-02, into the 3091 bp vector backbone of pLEH-02 which had been double digested with HincII and BamHI, generating pLEH-03 (Fig. 20). Restriction digestion was used to verify that the insert in pLEH-03 was oriented so as to restore both the HincII and BamHI sites, and that an insert of the expected size had been made. Finally, the 938 bp portion of the LS ORF coding for the C-terminus of LS was subcloned from pRO-02 which had been BamHI/StuI-digested, into the 3218 bp vector backbone of BamHI/SmaI-digested pLEH-03, generating pLEH-04 (Fig. 21). Restriction digests were used to confirm the size of the insert in pLEH-04, and that the insert had been made so as to reconstruct the BamHI site within the preS2 region.

To verify the DNA sequences of the subcloned S, MS, and LS regions, the dideoxy method (*e.g.*, Sequenase kit, United States Biochemical Corp.; Tabor and Richardson, 1989, J. Biol. Chem. 214:6447-6458) was used to sequence double stranded DNA from pT7T3/S, pT7T3/S2, and pLEH-04 inwards from primers (M13 Single Strand Primer, Pharmacia LKB cat. no. 27-1534-03, and M13 Reverse Sequence Primer, Cat no. 27-1532-01) complementary to both ends of the polylinker. In addition, a primer hybridizing just 3' of the ATG of the S ORF was used to facilitate sequencing of the 5' ends of the MS- and LS-encoding sequences. In order to sequence the interior of the S coding region, a 530 bp internal BalI/DraI fragment of the S region was

subcloned into pT7T318 which had been linearized with HincII. Sequence data generated using this strategy was analyzed in comparison to the published sequence of one isolate of subtype adw DNA (Ono et al., 1983, Nucl. Acids Res. 11:1747-1757). When we performed these experiments, our sequence data showed three translatable open reading frames corresponding to DNA encoding LS, MS and S. The nucleotide and deduced amino acid sequences of the LS (SEQ ID NO: 5 and NO: 6), MS (SEQ ID NO: 7 and NO: 8), and S (SEQ ID NO: 9 and NO: 10) regions are shown in Figures 22, 23, and 24, respectively. Changes in nucleotide and amino acid sequence from the published sequence are noted in Table 4 and are believed to result from variability within the adw subtype.

TABLE 4

Comparison of Published HBV Surface ORFs (subtype adw)* and Surface ORFs in pLEH-04

Base of Change Relative to A of ATG	Region of Change	Published Triplet	Published Amino Acid	pLEH-04 Triplet	pLEH-04 Amino Acid
27	preS1	A A C	Asn	A A T	Asn
40	preS1	C T T	Leu	T T T	Phe
83, 84	preS1	A C A	Thr	A A C	Asn
166	preS1	C T C	Leu	T T C	Phe
187	preS1	A T T	Ile	G T T	Val
336	preS2	A A T	Asn	A A C	Asn
348	preS2	T T G	Leu	T T C	Phe
388	preS2	C T T	Leu	T T T	Arg
488	preS2	A T C	Ile	A A C	Asn
887	S	A A G	Lys	A T G	Met
954	S	T C C	Ser	T C T	Ser

Ono, T. et al., Nucleic Acids Research 11:1747-1757 (1983).

6.1.2. CORE ANTIGENS

Hepatitis B virus core antigen-encoding sequences were also subcloned and engineered so as to be transcriptionally controlled by a vaccinia or vaccinia-like promoter. Like the surface antigen constructs, the hepatitis B core antigen constructs were designed so as to maintain a close proximity between the vaccinia promoter and the translational initiation codon of the core open reading frame. A gene encoding a full-length core was used. In addition, a deleted version of the core gene, referred to as coreΔ8, in which 8 amino acids are deleted as set forth below starting 16 amino acids from the 3' end of the full-length core antigen, was used.

CoreΔ8:

pAM6 was digested with StyI at two sites, one 16 bp upstream of the initiating ATG of core and the second 5 bp downstream of the C-terminal end of the core antigen coding region, and the overhanging single-stranded ends were blunted by treatment with Klenow fragment of DNA polymerase I by standard protocols. This 584 bp fragment was then gel-purified and subcloned into HincII-linearized pT7T318, generating a new plasmid, named pRO-09B (Fig. 25), containing three BglII sites. The orientation of the insert in pRO-09B was determined to be such that the EcoRI site in the polylinker of pT7T318 is located at the C-terminus of the open reading frame; digestion with BglII and EcoRI yielded a restriction map consistent with fragments of 24, 68, 420, and 2954 bp.

The 24 bp coreΔ8 deletion was generated using a partial BglII digest. By adjusting reaction conditions, it was possible to limit the action of BglII such that some of the product of the reaction was cut only at one of the three BglII sites in core.

This linearized plasmid, which consisted of a mixture of the pRO-09B plasmid which had been linearized at BglII site #1, BglII site #2, and BglII site #3, was then gel- purified. This linear product was then digested to completion with SmaI at the 3' end of the polylinker, and the large 3390 bp and 3414 bp vector fragments corresponding to initial linearization at either BglII #2 or BglII #3 were gel-purified. The 158 bp BglII/HincII fragment of pAM6 containing the C-terminus of the core ORF was then subcloned into the purified vector mixture. The resultant minipreps were screened for the presence of a 420 bp band upon BglII digestion, indicative of insertion in the correct orientation in which the BglII site has been recreated. The core regions of a number of candidate minipreps were sequenced to identify a clone in which BglII #2 was the site of digestion in the initial partial digest. The resulting plasmid, called pT7T3/C (Fig. 26), contained the expected 24 bp BglII-BglII deletion from bases 508 to 531 of the core ORF, and also had an unexpected single base pair deletion of an adenine residue at base 237 of 555 total base pairs within the full-length core open reading frame in comparison with the published sequence (Ono et al., *supra*). Since this deletion would result in a frame-shift upon translation of the core RNA, a method of reinserting this base into the open reading frame was used, as described below. Besides our intentional 24 bp deletion and this unintentional deletion at base 237, we found 11 additional nucleotides which differed from the published wild-type core sequence, which are summarized in Table 5. Only one of these nucleotide changes results in a change in amino acid sequence.

TABLE 5

Comparison of Published HBV Core (subtype adw)* and Core in pCODM/24

Base of Change Relative to A of ATG	Published Triplet	Published Amino Acid	pCODM/24 Triplet	pCODM/24 Amino Acid
81	G T A	Val	G T C	Val
82	C G A	Arg	A G A	Arg
117	C G A	Arg	C G G	Arg
156	C A T	His	C A C	His
159	A C T	Thr	A C C	Thr
229, 231	C A A	Gln	G A G	Glu
243	T C C	Ser	T C A	Ser
246	A G A	Arg	A G G	Arg
288	A A G	Lys	A A A	Lys
291	A T C	Ile	A T T	Ile
360	G T C	Val	G T A	Val

* Ono, T. et al., Nucleic Acids Research 11:1747-1757 (1983).

To reinsert the missing adenine at base 237 of the core ORF, pT7T3/C was subjected to oligonucleotide-directed site-specific mutagenesis on double-stranded DNA by published techniques (see for example Inouye, S. and Inouye, M., "Oligonucleotide-directed Site-specific Mutagenesis Using Double-Stranded Plasmid DNA", in *Synthesis and Applications of DNA and RNA*, ed. S. Narang, Academic Press). (Polymerase chain reaction could also be used to reinsert the missing nucleotide.) An oligonucleotide of sequence 5' GGA GGA TCC aGC ATC AAG G 3' (SEQ ID NO:11) was used both to replace the deleted base pair (lower case "a") and to generate for cloning purposes, without changing the core protein's amino acid sequence, a BamHI site (underlined) several bases 5' of the missing base pair. Mutation using

this procedure generated two types of DNA plasmids: wild-type and mutant containing the new BamHI site and the inserted adenine. The mutant DNA was then cloned selectively using this BamHI site, such that each of the two pieces of DNA on either side of the BamHI site were first cloned individually and then reunited. To do this, the transformed *E. coli* containing a mixture of mutated and non-mutated plasmids were inoculated into 250 ml LB broth with 50 μg/ml ampicillin, and a purified large-scale preparation of the DNA was made by standard techniques (Sambrook et al., *supra*). Separate aliquots of this DNA were then digested with either EcoRI/BamHI or HindIII/BamHI. Part of the EcoRI/BamHI-digested aliquot was subjected to electrophoresis on an agarose gel, and a 448 bp fragment was excised and purified. The same method was used to purify the 269 bp HindIII/BamHI fragment. Use of this method generated two fragments which were then used to make two plasmids: pT7T3/CHB (Fig. 27), in which the 269 bp 5' HindIII to BamHI region is cloned into the 2853 bp vector fragment of HindIII/BamHI digested pT7T318, and pT7T3/CRB (Fig. 28), in which the 448 bp 3' BamHI to EcoRI region is cloned into the 2862 bp vector fragment of BamHI/EcoRI-digested pT7T318. In both cases, the inserts were verified by restriction digestion to be oriented so as to recreate both restriction sites.

To reunite the core open reading frame, the 448 bp BamHI/EcoRI fragment of pT7T3/CRB containing the 3' end of the mutated open reading frame was then cloned into the 3101 bp fragment of BamHI/EcoRI-digested pT7T3/CHB so as to preserve both the BamHI and EcoRI sites, generating pT7T3/CODM (Fig. 29), which contains a complete coreΔ8 gene containing the desired single base pair insertion. The nucleotide

sequence (SEQ ID NO:12) and deduced amino acid sequence (SEQ ID NO:13) of the coreΔ8 ORF in pT7T3/CODM are shown in Figure 30.

Full-length core:

In order to construct a full-length core gene, pT7T3/CRB was digested with BglII, which linearizes the plasmid, and ligated to a pair of oligonucleotides designed to hybridize to each other, add 24 bp of sequence, and leave overhanging BglII-compatible ends. Oligonucleotides of sequence 5' GAT CCC AAT CGC GGC GTC GCA GAC 3' (SEQ ID NO:16) and 5' GAT CGT CTG CGA CGC GGC GAT TGG 3' (SEQ ID NO:17), were used for this purpose, and were designed to make changes in the nucleotide sequence at each end of the insert so as to destroy both BglII sites but to still encode the correct amino acid sequence. By digesting the religated plasmid with BglII, selection was imposed for an insertion mutant, pCRB/24 (Fig. 31), which lacks BglII sites. pCRB/24 is similar to pT7T3/CRB except that pCRB/24 encodes 8 additional amino acids of core antigen sequence and lacks any BglII sites. The correct orientation of the insert to recreate the authentic core amino acid sequence, and the fact that only a single insert had been made, was confirmed by dideoxy sequencing of pCRB/24. To reconstruct the full-length core gene, pCRB/24 was digested with EcoRI and BamHI, and the 472 bp fragment containing the 3' end of the core gene was subcloned into the 3101 bp vector fragment of EcoRI/BamHI-digested pT7T3/CHB to generate plasmid pCODM/24 (Fig. 32; Table 5, *supra*). The orientation of the insert in pCODM/24 was such that both the BamHI and EcoRI sites were recreated. pCODM/24 is identical to pT7T3/CODM except that pCODM/24 contains 24 bp which are deleted in pT7T3/CODM. The nucleotide

sequence (SEQ ID NO:14) and deduced amino acid sequence (SEQ ID NO:15) of the core region in pCODM/24 are shown in Figure 33.

6.2. SINGLE EXPRESSION VECTORS FOR S, MS, LS AND CORE ANTIGENS

A variety of expression vectors for insertion into the vaccinia thymidine kinase gene was constructed in which p7.5 (Fig. 2), p11 (Fig. 4), or the modified p7.5 promoter (Fig. 6) directed the expression of one of the cloned HBV genes. A number of constructs were made in which expression of a single HBV antigen was directed by the p7.5 promoter. To facilitate this, a unique XhoI restriction site was created downstream of the p7.5 promoter in pGS53. This was done by linearizing pGS53 with BamHI, blunting the ends using Klenow treatment, setting up a ligation reaction in the presence of an excess of an XhoI linker of sequence 5' CCC TCG AGG G 3' (SEQ ID NO:18) (New England Biolabs, cat. no. 1072), heating to denature, and religating. A fraction of this reaction was used to transform *E. coli* (DH5α $F'I^Q$), and minipreps of colonies were screened for the presence of a new XhoI site. The plasmid thus created, pGS53X (Fig. 34), contains two BamHI sites, one on each side of the XhoI site.

The 824 bp S fragment was then excised from pT7T3/S by digestion with HindIII and KpnI followed by treatment with T4 DNA polymerase in the presence of deoxyribonucleotides. This treatment served to fill in overhanging 5' ends and remove overhanging 3' ends. The 1251 bp MS fragment was excised from pT7T3/S2 by digestion with KpnI and SphI followed again by T4 polymerase treatment. The 1325 bp LS fragment was excised from pLEH-04 by digestion with HindIII and EcoRI followed by filling in the overhanging 5' ends

with the Klenow fragment of DNA polymerase I. The ends of all three gel-purified fragments thus generated were made XhoI-compatible by addition of the XhoI linkers described above, which in this case had been phosphorylated using T4 polynucleotide kinase and standard procedures, followed by digestion with XhoI. Each fragment was then subcloned into the 4704 bp XhoI-linearized and BAP-dephosphorylated pGS53X. The single antigen expression vectors thus created contained the S gene, pGS53/S (Fig. 35), the MS gene, pGS53/S2 (Fig. 36), or the LS gene, pGS53/S1 (Fig. 37), each under the control of the p7.5 promoter. The orientation of each insert was determined using conventional restriction mapping to be such that each open reading frame could be expressed from the p7.5 promoter. Specifically, the correct orientation of the insert was confirmed by HincII digestion to yield restriction maps consistent with fragments from pGS53/S of 352 and 5186 bp, from pGS53/S2 of 514 and 5447 bp, and from pGS53/S1 of 321, 520 and 5202 bp. The genes in these plasmids were engineered so that the 3' end of the p7.5 promoter in these vectors is within approximately 20 bp of the initiating ATG. In each of these constructs, the expression cassette region is flanked by sequence from the vaccinia virus thymidine kinase gene. This thymidine kinase gene is contained in the HindIII-J fragment of the vaccinia genome.

A similar series of constructs for expression of S, MS, and LS antigens individually directed by the modified p7.5 promoter and designed for insertion into the vaccinia TK gene was made as follows. The S region from pT7T3/S and the LS region from pLEH-04 were amplified by polymerase chain reaction using techniques well known in the art

(Erlich, H.A., ed., *PCR Technology, Principles and Applications for DNA Amplification*, M. Stockton Press, New York, 1989). One primer used was of sequence 5' CTC ACT CGA GGG AAT AAG CT 3' (SEQ ID NO:19), which hybridizes from bases -23 to -4 relative to the initiating ATG in both plasmids. This primer contains a 2 bp change from the plasmid sequence, AA to CG at positions 7 and 8, which was designed to create an XhoI site (CTCGAG) upon amplification. Another primer used, 5' TTG TTA GGC CTT AAA TGT AT 3' (SEQ ID NO:20), hybridizes at the carboxy-terminal end of the open reading frames in both of these plasmids (underlined nucleotides indicate position of termination codon TAA on the complementary strand). This primer contains a 2 bp change from the plasmid sequence, GT to CC at positions 9 and 10, designed to create a StuI site (AGGCCT) upon amplification. The fragments generated by PCR of pT7T3/S and pLEH-04 with these primers are expected to be of sizes 714 and 1203 base pairs, respectively. These pieces were then digested with XhoI and StuI to remove the small fragments at either end, yielding fragments of size 699 and 1188 base pairs, respectively, which were purified from an agarose gel. These latter fragments were inserted into the 4026 bp vector fragment of XhoI and StuI double digested pSC59 so as to reconstruct XhoI and StuI sites at each end of the insert. The S-containing construct named pHTL-8 (Fig. 38), and the LS containing construct, pHTL-9 (Fig. 39), each consist of the modified p7.5 promoter situated so as to direct expression of the HBV antigen located downstream.

A construct designed to allow expression of the MS antigen from the modified p7.5 promoter was made as follows. A 998 bp EcoRI-StuI fragment of

pT7T3/S2 was gel-purified and inserted into the 4032 bp EcoRI-StuI digested vector fragment of pSC59 so as to reconstruct the EcoRI and StuI sites at either end of the insert. The resulting expression plasmid, pHTL-10 (Fig. 40), contains the MS ORF located downstream of the modified p7.5 promoter. pHTL-10 contains the sequence "CCCATG" surrounding the ATG of preS2. According to M. Kozak (1989, J. Cell Biol. 108:229-241), the DNA sequence leading to the strongest ribosome binding to a translation initiation codon is "ANNATG", where N is any nucleotide. In the absence of a strong consensus-like sequence surrounding the ATG of preS2, significant levels of expression of the S antigen from the MS-containing constructs may be observed by western blot. Polymerase chain reaction may be used to modify the surroundings of the preS2 ATG in pHTL-10 or other constructs to match this consensus sequence, as follows, with the goal of shifting expression from S to MS.

A fragment of pHTL-10 containing the initiating ATG of preS2 was amplified by PCR, and by the proper choice of primers the sequence upstream of the ATG was made to match the consensus sequence described by Kozak. One primer used was of the sequence 5' GAG CTC GGT ACC ACC ATG AAG TGG A 3' (SEQ ID NO:21), which hybridizes from bases -15 to +10 relative to the preS2 ATG and contains a KpnI site (GGTACC) as well as the desired C to A change at position 13. A second primer used in conjunction with this first primer is of sequence 5' GAG CAG GGG TCC TAG GAA TC 3' (SEQ ID NO:22). This second primer hybridizes to the opposite strand from the first primer, from bases +204 to +185 relative to the preS2 ATG in pHTL-10, and contains an AvrII site (CCTAGG).

The product of a PCR reaction using these primers was digested with KpnI and AvrII, run on an agarose gel, and the 197 bp fragment containing the initiating ATG of MS was gel-purified. This fragment was cloned into the 4833 bp gel-purified vector fragment of pHTL-10 generated by digestion with KpnI and AvrII so as to retain both restriction sites. Dideoxy sequencing was used to confirm that the expected changes had been made upstream of the initiating ATG of preS2 in the new plasmid, pHTL-10M (Fig. 41), to match the consensus sequence. Western blotting was used to compare the relative expression levels of MS vs. S in pHTL-10 and pHTL-10M (see Section 8.2 *infra*). Analysis by Western blot probed with an anti-S polyclonal antibody of proteins expressed transiently from pHTL-10 and pHTL-10M in vaccinia-infected cells showed that the amount of MS protein expressed from pHTL-10M was significantly increased and the amount of S protein significantly decreased relative to that expressed from pHTL-10.

6.3. SINGLE EXPRESSION VECTORS FOR CORE-S-REGION FUSIONS

A series of expression constructs consisting of fusions of the core antigen to the preS2 region or to part of the S or preS1 regions was also made. The fusion polypeptides were expressed from the modified p7.5 promoter (Fig. 6). The first step in generating core to S-region fusions was to subclone the full-length core gene downstream of this promoter. This was done by digesting pCODM/24 with PstI, treating with mung bean nuclease (New England Biolabs) to remove single-stranded overhangs according to the manufacturer's directions, and then digesting with EcoRI. The 727 bp fragment thus generated containing the wild-type core ORF was isolated and subcloned into

the 4030 bp vector backbone of pSC59 which had been digested with XhoI, treated with mung bean nuclease, and then digested with EcoRI. Restriction digestion was used to verify that the insert was oriented so as to retain the EcoRI site at the C-terminal end of the core antigen. The resulting plasmid, pHTL-25 (Fig. 42), contains the full-length core antigen oriented so as to permit expression of the core from the modified p7.5 promoter. The expression cassette of promoter and antigen in pHTL-25 is flanked by pieces of the vaccinia TK gene to allow *in vivo* recombination of the plasmid into the vaccinia TK gene.

A vector was constructed to allow expression of a core-preS1 fusion polypeptide (called core-preS1*, because only a portion of preS1 is present). This core-preS1* fusion consists, starting at the amino terminus of the fusion protein, of amino acid residues 1 through 145 of core, a 3 amino acid spacer (ser-ala-cys), amino acids 1 through 56 of preS1, and a 4 amino acid tail (arg-pro-thr-ser; SEQ ID NO:56). To make this fusion vector, a piece of the preS1 region corresponding to the first 56 amino acids of preS1 was amplified by polymerase chain reaction from pLEH-04. We expected this region of preS1 to be immunogenic; data presented by Gassin et al. showed that the anti-preS1 immune responses of health workers vaccinated with a plasma-derived HBV vaccine were primarily against amino acid residues 12-27 and 32-53 (Gassin et al., 1990, Prevalence of preS1 peptides antibodies following HEVAC B vaccine, in *Progress in Hepatitis B Immunization*, Coursaget, P. and M.J. Tong, eds., John Libbey Eurotext, London). One primer used was of sequence 5' TCA CTA TCC GGA ATC AGC TT 3' (SEQ ID NO:23), which hybridizes from bases -22 to -3 relative to the ATG of preS1 and is designed

to introduce a BspEI site (TCCGGA) upon amplification and to destroy a potential termination codon (underlined C was changed from an A in pLEH-04). A second primer used was of sequence 5' GGT GAA TTC TGG CCC GAA TG 3' (SEQ ID NO:24), which hybridizes to the opposite strand from bases +171 to +152 of the preS1 region and is designed to create an EcoRI site (GAATTC) upon amplification. The product of the PCR reaction was double digested with BspEI and EcoRI, and the 178 bp product was gel- purified and cloned into the 4483 bp BspEI and EcoRI double digested vector fragment of pHTL-25 in an orientation which recreated both the EcoRI and BspEI sites at the ends of the insert. The resulting plasmid, pHTL-26 (Fig. 43), consists of the modified p7.5 promoter oriented so as to drive expression of the core-preS1* fusion. The nucleotide sequence (SEQ ID NO:25) and predicted amino acid sequence (SEQ ID NO:26) of the core-preS1* fusion made from pHTL-26 are shown in Figure 44.

Using a similar design to that of pHTL-26, a construct was made for expression of a fusion consisting, starting at the amino terminus of the fusion, of amino acid residues 1 through 144 of core, a one amino acid spacer (asp), residues 1 through 55 of preS2, residues 1 through 8 of S, and a 4 amino acid tail (arg-pro-thr-ser; SEQ ID NO:56). The preS2- and S-region insert in this construct was generated by polymerase chain reaction amplification of pLEH-04. One primer used was of sequence 5' CAT CCT CCG GAC ATG CAG TG 3' (SEQ ID NO:27), hybridizes from base +313 of the preS1 region to base +8 of preS2, and was designed to create a new BspEI site (TCCGGA) during amplification. A second primer which was used in conjunction with this first primer was of sequence 5' GTC CTA GGA ATT CTG ATG TG 3' (SEQ ID NO:28), which

hybridizes to the opposite strand from the first from bases +31 to +12 of the S region, and creates a new EcoRI site (GAATTC) during amplification. The product of PCR of pLEH-04 using these primers was digested with EcoRI and BspEI, and the 190 bp fragment was gel-purified. This fragment was then cloned in a manner so as to regenerate EcoRI and BspEI sites at either end of the insert into the 4483 bp gel-purified vector fragment of pHTL-25 which had been double digested with EcoRI and BspEI. The resulting plasmid, pHTL-27 (Fig. 45), is designed to allow expression of a core to preS2 fusion from the modified p7.5 promoter. The nucleotide sequence (SEQ ID NO:29) and predicted amino acid sequence (SEQ ID NO:30) of the core-preS2 fusion made from pHTL-27 are shown in Figure 46.

A third construct of this type, consisting from the amino terminus of amino acid residues 1 through 144 of core, a one amino acid spacer (asp), residues 107 through 163 of the S antigen, and a 7 amino acid tail (asn-ser-gly-leu-leu-val-lys; SEQ ID NO:57) was made in a similar manner. This fusion polypeptide is referred to herein as core-S* because only a small portion of the S antigen, judged to be most important in terms of antigenicity, is present. Hydrophilic stretches of the S antigen expected to contain antigenic determinants are located between residues 110 and 156 of S (Bhatnagar et al., 1982, Proc. Natl. Acad. Sci. USA 79:4400-4409). The nonapeptide sequence from 139 to 147 represents all or an essential part of the a determinant of HBsAg, the major epitope in S against which humans make antibodies after exposure to any subtype of the hepatitis B virus (id.). A 197 bp piece of the S region containing these immunogenic regions was generated by PCR amplification of pLEH-04 using

primers 5' GTA TGT TTC CGG ATT GTC CT 3' (SEQ ID NO:31), which hybridizes from bases +304 to +323 of the S region and was designed to generate a BspEI site (TCCGGA) upon amplification, and 5' TGA GGC CGA ATT CCA TAG GT 3' (SEQ ID NO:32), which hybridizes to the opposite strand from bases +500 to +481 of S and was designed to create an EcoRI site (GAATTC) upon amplification. The product of PCR on pLEH-04 using these primers was double digested with EcoRI and BspEI, and the 172 bp product containing the S fragment was gel-purified. This fragment was then inserted, in a manner such that the restriction sites were regenerated at either end of the insert, into the 4483 bp vector fragment of pHTL-25 which had also been doubly digested with EcoRI and BspEI. The resulting plasmid was called pHTL-28 (Fig. 47). The nucleotide sequence (SEQ ID NO:33) and predicted amino acid sequence (SEQ ID NO:34) of the core-S* fusion made from pHTL-28 are shown in Figure 48.

6.4. DUAL EXPRESSION VECTORS

A number of constructs were made consisting of two promoters, divergently oriented and subcloned so as to be located within the vaccinia TK gene. Various antigens were then subcloned downstream of one or both of the promoters, with the goal of allowing expression of two non-fused antigens from a single vaccinia recombinant containing the integrated plasmid. The first step in one scheme for making these constructs was to delete the β-galactosidase gene from pSC10 by digesting the plasmid with BamHI and religating the larger fragment. The resulting plasmid, pSC10ΔlacZ (Fig. 49), contains the 11 kD promoter located upstream of a unique BamHI site. At this point, the p7.5 and p11 promoters were placed

"back to back" to allow for future construction of a single expression vector capable of expressing two different antigens. To do this, the 117 bp XbaI-BamHI fragment of pSC10ΔlacZ was excised, blunted by Klenow treatment, and inserted into the 4690 bp fragment of pGS53 which had been linearized by HincII. The orientation of the insert in this new construct, pDPV (Fig. 50), was determined to be such that the two promoters are oriented back-to-back; double digestion with EcoRI and BamHI yielded a restriction map consistent with fragments of sizes 9, 35, 375 and 4392 bp. This process retains the XbaI site between the two promoters but destroys the BamHI site downstream of p11.

pDPV was modified to insert S downstream of the 11 kD promoter. This was done by cutting pDPV with XhoI and inserting the 834 bp S-containing XhoI fragment of pGS53/S into this 4811 bp linearized vector. Digestion of candidate miniprep DNAs with XbaI to yield restriction maps consistent with fragments of sizes 218, 1387, and 4040 bp was used to verify the correct orientation of the inserted S piece downstream of the p11 promoter in the resulting construct, pDPV-01 (Fig. 51). This construct contains an extra ATG 23 bp upstream of the S antigen initiating ATG which, if left in place, would interfere with proper expression of S from its own ATG. Oligonucleotide-directed mutagenesis was used to change this ATG to an ATA. One oligonucleotide used hybridizes to the strand containing the ATG from bases -12 to -30 relative to the initiating ATG of S, and has the sequence 5' CGA GGA ATT TAT TTA TAG C 3' (SEQ ID NO:35), where the underlined T indicates the base to be mutated. The sequence of the region upstream of

the S ATG in the resulting plasmid, pRO-10 (Fig. 52), was confirmed by dideoxy sequencing of the plasmid.

A vector, pRO-16 (Fig. 53), was constructed from pRO-10, allowing expression of the S antigen from the p11 promoter and the coreΔ8 antigen from the p7.5 promoter. This was done by cutting pT7T3/CODM with HindIII and EcoRI, filling in the overhanging 5' ends using the Klenow fragment of DNA polymerase I, and gel-purifying the 721 bp fragment containing the coreΔ8 antigen. This fragment was then subcloned into the 5645 bp vector fragment of SmaI-linearized pRO-10. The orientation of the insert in the resulting plasmid, pRO-16, was determined to be such that the two promoters were oriented back-to-back; BamHI digestion yielded a restriction map consistent with fragments of 275 and 6091 bp.

pRO-10 was modified to insert the MS antigen downstream of p7.5 as follows. The 1257 bp MS fragment was removed from pT7T3/S2 by double digestion with KpnI and HindIII, blunting the ends of the fragment with T4 DNA polymerase, and gel-purifying the 1.2 kb fragment. This piece was then inserted into SmaI-linearized, bacterial alkaline phosphatase (BAP)-treated pRO-10. The orientation of the insert in the resulting plasmid, pRO-11 (Fig. 54), was verified to be such that MS could be expressed from the p7.5 promoter; digestion with XbaI yielded fragments of sizes 218, 538, 1387 and 4759 bp. pRO-11 contains the S and MS ORFs downstream of p11 and p7.5, respectively.

The S fragment of pRO-11 was also replaced by coreΔ8. This was done by cutting pRO-11 with XhoI, blunting the ends of the pieces by Klenow treating, BAP-treating, and gel-purifying the vector fragment of 6072 bp. Into this blunt-ended vector was inserted

the coreΔ8 fragment of pT7T3/CODM, generated by cutting pT7T3/CODM with HindIII and EcoRI, blunting the ends with Klenow, and gel-purifying the 721 bp core-containing region. Orientation of the insert in the resulting plasmid, pRO-17 (Fig. 55), was determined to be such that coreΔ8 could be expressed from the p11 promoter; digestion with BamHI yielded a restriction map consistent with fragments of sizes 49, 654 and 6090 bp. pRO-17 contains coreΔ8 downstream of the 11 kD promoter and MS downstream of the 7.5 kD promoter.

A plasmid analogous to pRO-17 but containing the full-length core gene instead of coreΔ8 was made in a manner identical to pRO-17, except that the 745 bp insert was cut from pCODM/24 instead of pT7T3/CODM. The same vector and restriction sites were used. The orientation of the insert in the resulting plasmid, pRO-22 (Fig. 56) was determined to be such that full-length core should be expressed from the 11 kD promoter; BamHI digestion yielded a restriction map consistent with fragments of sizes 49, 654 and 6114 bp.

A vector was also created containing only the divergently oriented p7.5 and p11 promoters by cutting pRO-10 with XhoI to excise the 834 bp S-containing fragment and religating the large 4811 bp vector fragment to delete the S region. The resulting plasmid, pRO-12 (Fig. 57), contains only the p7.5 and p11 promoters and no HBV antigens. This plasmid differs from pDPV in that it lacks the extra ATG downstream of the p11 promoter which was removed in the course of making pRO-10 from pDPV-01.

From pRO-12, a vector was created consisting of the LS antigen inserted downstream of p7.5. This construct was made by cutting the LS antigen from

pLEH-04 using a HindIII and EcoRI double digest, filling in the overhanging 5' ends by Klenow treatment, and gel-purifying the 1329 bp fragment. This LS fragment was then subcloned into the 4811 bp vector fragment of SmaI-digested, BAP-treated pRO-12. The orientation of the insert in the resulting plasmid, pRO-13B (Fig. 58), was determined to be such that LS should be expressible from p7.5; digestion with XbaI and PstI yielded a restriction map consistent with fragments of 125, 224, 641, 646, and 4504 bp. pRO-13B therefore consists of the p11 and p7.5 promoters oriented back-to-back with the LS open reading frame positioned downstream of p7.5.

The coreΔ8 antigen was inserted downstream of p11 in pRO-13B. This was done by digesting pT7T3/CODM with HindIII and EcoRI, filling in the overhanging 5' ends with the Klenow fragment, and gel-purifying the 721 bp coreΔ8-containing fragment. This fragment was then subcloned into the 6144 bp XhoI-linearized vector pRO-13B which had been Klenow- and then BAP-treated. The orientation of the insert in the resulting plasmid, pRO-18 (Fig. 59), was determined to be such that the coreΔ8 open reading frame should be expressed from the 11 kD promoter; digestion with BamHI yielded a restriction map consistent with fragments of sizes 376, 654 and 5835 bp.

A vector featuring coreΔ8 downstream of p11 was made by cutting pRO-17 with XbaI to remove the 1496 bp region containing the p11 promoter, core, and most of one-half of the vaccinia DNA, and cloning this piece into the 4040 bp XbaI-cut, BAP-treated vector fragment of pRO-12. Orientation of the insert in the resulting plasmid, pRO-19 (Fig. 60), was determined to be such that the two promoters, p7.5 alone and p11

upstream of coreΔ8, were divergently oriented; digestion with BamHI yielded a restriction map consistent with fragments of sizes 654 and 4882 bp.

A plasmid analogous to pRO-19 containing the p11-full-length core antigen cassette was made by cutting pCODM/24 with HindIII and EcoRI, blunting the ends by Klenow treatment, and gel-purifying the 745 bp fragment. This core-containing fragment was subcloned into the 4811 bp XhoI-linearized pRO-12 which had also been made blunt-ended by Klenow. The resulting plasmid features full-length core downstream of p11. Orientation of the insert in the resulting plasmid, pRO-19/24 (Fig. 61), was determined to be such that the full-length core gene could be expressed from the 11 kD promoter; digestion with BamHI yielded a restriction map consistent with fragments of sizes 654 and 4906 bp.

To allow production of a series of constructs featuring the modified p7.5 promoter and the p11 promoter back to back, another vector was made. pRO-19/24 was double digested with XbaI and EcoRI to excise the 893 bp fragment containing the p11-full-length core expression cassette. This fragment was gel-purified, the 5' overhanging ends filled in by Klenow fragment of DNA polymerase I, and the piece subcloned into the 4042 bp vector fragment of pSC59 which had been digested with HindIII and treated with Klenow and BAP. To determine that the insertion had been made such that the two promoters, p11 and the modified p7.5, were divergently oriented, the resulting plasmid, pHTL-5 (Fig. 62), was digested with PvuII and HindIII to yield a restriction pattern consistent with fragments of 1735 and 3200 bp. pHTL-5 contains the full-length core gene expressed from the

p11 promoter, and the modified p7.5 promoter without any antigen cloned downstream.

From pHTL-5, plasmids expressing S or LS from the modified p7.5 promoter and full-length core from the 11 kD promoter were made as follows. The S region from pT7T3/S and the LS region from pLEH-04 were removed by polymerase chain reaction using the same pair of primers. One primer used was of sequence 5' CTC ACT CGA GGG AAT AAG CT 3' (SEQ ID NO:36), and was designed to hybridize from bases -23 to -4 relative to the initiating ATG in each plasmid and to generate an XhoI site (CTCGAG) upon amplification. A second primer used in conjunction with the first primer was of sequence 5' TTG TTA GGC CTT AAA TGT AT 3' (SEQ ID NO:37), and was designed to hybridize to the opposite strand from the first primer from bases +1179 to +1160 relative to the ATG of LS in pLEH-04 and from bases +690 to +671 relative to the ATG of S in pT7T3/S, and to generate a StuI site (AGGCCT) upon amplification. The fragments generated by PCR using these primers on pT7T3/S and pLEH-04 were of sizes 714 and 1203 base pairs, respectively. These amplified fragments were digested with XhoI and StuI, removing a few base pairs at either end and yielding fragments of size 699 and 1188 base pairs, respectively, which were purified from an agarose gel. These fragments were cloned so as to regenerate XhoI and StuI sites at either end of the insert, into the 4947 bp vector fragment of pHTL-5 which had been digested with XhoI and StuI. The S- and LS-containing plasmids generated in this manner were called pHTL-6 (Fig. 63) and pHTL-7 (Fig. 64), respectively.

A similar construct consisting of MS expressed from the modified p7.5 promoter and the full-length core antigen expressed from the 11 kD

promoter were made as follows. The 998 bp MS-containing fragment of EcoRI-StuI digested pT7T3/S2 was gel-purified and subcloned into the 4929 bp EcoRI-StuI vector fragment of pHTL-5. Restriction mapping was used to verify that the insertion had been made in an orientation so as to regenerate the EcoRI site at the 5' end encoding the amino terminus of the MS insert and the StuI site at the 3' end encoding the carboxyl terminus in the resulting plasmid, pHTL-11 (Fig. 65).

A plasmid for expressing MS from the modified p7.5 promoter and coreΔ8 from the p11 promoter was made from pHTL-11. To do this, a 849 bp AvrII-BspEI fragment of pHTL-11 containing the two promoters and the 5' end of each open reading frame was gel-purified and cloned into the 5742 bp AvrII-BspEI vector fragment of pRO-17 containing the 3' ends of MS and coreΔ8. The insert was made in such a way as to regenerate the AvrII and BspEI sites, and the resulting plasmid was called pHTL-12 (Fig. 66). In this plasmid, the coreΔ8 antigen has replaced the wild-type core antigen which was present in pHTL-11.

From pHTL-12 was made another construct, pRO-16M (Fig. 67), which consists of the coreΔ8 antigen cloned downstream of the p7.5 promoter. This construct was planned so as to delete an out-of-frame ATG which is present 31 bp upstream of the correct initiating ATG of coreΔ8 in pRO-16. pRO-16M was made by linearizing pRO-12 with BamHI, which cuts downstream of the p7.5 promoter, blunting the ends with Klenow treatment, and then treating with BAP. The 719 bp coreΔ8-containing insert was gel-purified from pHTL-12 which had been digested with PstI and then treated with mung bean nuclease. From the ligation of this fragment of pHTL-12 into the

linearized pRO-12 were chosen clones containing the coreΔ8 antigen situated downstream of p7.5. The presence and correct orientation of the insert in pRO-16M were determined by PstI/BspEI digestion, which yielded a restriction pattern consistent with fragments of 850 and 4684 bp.

From pRO-16M was made an additional construct, pHTL-30 (Fig. 68), consisting of MS expressed from the modified p7.5 promoter and coreΔ8 expressed from the p7.5 promoter. The p7.5-coreΔ8 cassette was excised from pRO-16M by PCR. One primer used for this reaction was of sequence 5' GTG GGT AAG CTT CTC GAT GT 3' (SEQ ID NO:38) and is designed to hybridize from bases -230 to bases -211 relative to the ATG of coreΔ8, upstream of the p7.5 promoter, and to create a HindIII site (AAGCTT) upon amplification. A second primer used in combination with the first primer was of sequence 5' AGT TTC CAA GCT TAT GAG 3' (SEQ ID NO:39), and was designed to hybridize to the opposite strand from bases +563 to +546 and to generate a HindIII site upon amplification. The product of a PCR reaction using these two primers was digested with HindIII and the 774 bp fragment containing p7.5-coreΔ8 was gel-purified. This fragment was inserted into pHTL-10M which had been linearized with HindIII just upstream of the modified p7.5 promoter and BAP treated. HindIII digestion was used to verify the presence of the insert in pHTL-30. Digestion with EcoRI/BspEI yielded a restriction map consistent with fragments of 706 and 5098 bp, verifying that the two promoters were oriented back-to-back.

Using pHTL-5, a variant of pRO-16 consisting of S expressed from the p11 promoter and full-length core, instead of coreΔ8, expressed from the p7.5

promoter was made. This was done by cutting pHTL-5 with BspEI and SacI to excise the 268 bp fragment coding for the carboxyl terminus of full-length core. This fragment was gel-purified and inserted into the 6123 bp vector fragment of pRO-16 which had been digested with BspEI and SacI and gel-purified. The orientation of the insert in the resulting plasmid, pHTL-13 (Fig. 69), was determined, based on the recreation of the BspEI and SacI sites, to be such that the two promoters are divergently oriented.

From pHTL-13 was made a plasmid, pHTL-15 (Fig. 70), which contains the S antigen cloned downstream of the modified p7.5 promoter and full-length core cloned downstream of the p7.5 promoter. This plasmid was made by digesting pHTL-13 with XbaI and SacI to excise the fragment containing the p7.5-full-length core cassette. This fragment was gel purified, treated with mung bean nuclease, and the 1008 bp resulting fragment inserted into the 4729 bp vector fragment of pHTL-8 which had been linearized with BclI, Klenow filled, and BAP-treated. Presence of and correct orientation of the insert in pHTL-15 was verified by digestion with PstI to yield a restriction map consistent with pieces of 2223 and 3514 bp. This plasmid may be used for additional cloning; however, it is preferably not used for making recombinant vaccinia viruses because there is a segment of about 600 bp of vaccinia DNA between the two expression cassettes which may interfere with the recombination of both cassettes into the vaccinia genome. However, by deletion of this sequence, a plasmid useful for making recombinant viruses can be made.

From pHTL-13 was made another plasmid, pHTL-17 (Fig. 71), containing MS downstream of the

modified p7.5 promoter and full-length core downstream of the p7.5 promoter. This plasmid was made by digesting pHTL-13 with XbaI and SacI to excise the fragment containing the p7.5-full-length core cassette. This fragment was gel-purified, treated with mung bean nuclease, and the 1008 bp resulting fragment inserted into the 5034 bp vector fragment of pHTL-10 which had been linearized with BclI and Klenow filled. Presence of the insert in pHTL-17 was verified by digestion with PstI, which yielded a restriction map consistent with pieces of sizes 1000, 2223, and 2819 bp. Orientation of the insert was verified by this digest to be such that the two promoters are back-to-back. This plasmid may be used for additional cloning; however, it is preferably not used for making recombinant vaccinia viruses because there is a segment of about 600 bp of vaccinia DNA between the two expression cassettes which may interfere with the recombination of both cassettes into the vaccinia genome. However, by deletion of this sequence, a plasmid useful for making recombinant viruses can be made.

From pHTL-15 and pHTL-30 was made another plasmid, pHTL-33 (Fig. 72), consisting of full-length core cloned downstream of p7.5 and MS with the upregulated ATG cloned downstream of the modified p7.5 promoter. This plasmid was made by polymerase chain reaction amplification of the p7.5-full-length core cassette out of pHTL-15 using primers of the sequence 5' CTC GAT GTC GAC TAG CCA TA 3' (SEQ ID NO:58), which hybridizes to pHTL-15 from positions -237 to -218 relative to the ATG of core and the last four bases of which are part of an NdeI site (CATATG), and 5' AAG TTG TCG ACC TTA TGA GT 3' (SEQ ID NO:59), which hybridizes to the opposite strand from positions +587

to +568, downstream of the termination codon for full-length core, and which contains a SalI site (GTCGAC). The product of the PCR reaction was digested with SalI and NdeI and the 798 bp fragment containing the core expression cassette was isolated from an agarose gel and cloned into the 5046 bp gel purified vector fragment of pHTL-30 which had been digested with SalI and NdeI. Presence of the insert and situation of the two cassettes in a back-to-back orientation in pHTL-33 was determined by digestion with SalI and NdeI to yield a restriction map consistent with fragments of sizes 798 and 5046 bp.

From pHTL-33 was made an additional plasmid, pHTL-34 (Fig. 73), consisting of S expressed from the modified p7.5 promoter and full-length core expressed from the p7.5 promoter. This plasmid was made by digesting pHTL-8 with XhoI and ClaI and gel purifying the 767 bp fragment containing the S antigen. This fragment was cloned into the gel purified 4772 bp vector fragment of pHTL-33 which had also been digested with XhoI/ClaI. Presence and correct orientation of the insert in pHTL-34 were verified by digestion of the plasmid with XhoI/ClaI to yield a restriction map consistent with fragments of 767 and 4772 bp.

Also from pHTL-33 was made another plasmid, pHTL-35 (Fig. 74), consisting of LS expressed from the modified p7.5 promoter and full-length core expressed from the p7.5 promoter. This plasmid was made by gel purifying the 1256 bp XhoI/ClaI fragment of pHTL-9 containing the LS antigen and inserting it into the 4772 bp vector fragment of pHTL-33 which had also been digested with XhoI/ClaI. Presence and correct orientation of the insert in pHTL-35 were verified by digestion of the plasmid with XhoI and ClaI to yield a

restriction map consistent with fragments of 1256 and 4772 bp.

6.5. PRES1-CORE FUSION CONSTRUCTS

A series of constructs consisting of preS1-core fusions was also made. PreS1-core fusions included only the preS1 region of the LS antigen (*i.e.*, lacking the preS2 and S regions) fused to one of the two types of core (full length or coreΔ8). pPB-05 (Fig. 76) contains the preS1-encoding region fused to the coreΔ8 gene, with preS1 constituting the amino terminus of the resulting fusion protein. pPB-09 (Fig. 77) contains the preS1-encoding region fused to the full-length core gene in the same orientation. Both of these constructs consist of inserts into the standard cloning vector, pT7T318, so that the open reading frames are flanked by restriction endonuclease sites and can be further manipulated into expression vectors as needed.

To make pPB-05, the coreΔ8 open reading frame was cloned by polymerase chain reaction amplification from pT7T3/CODM. The primers used to accomplish this were of sequence 5' GCG CCA TGG ACA TTG ACC CTT ATA 3' (SEQ ID NO:40), which hybridizes from bases -5 to +19 relative to the initiating ATG of coreΔ8, and 5' CCC TGA TCA CTA ACA TTG AGA TTC CCG A 3' (SEQ ID NO:41), which hybridizes to the opposite strand from bases +543 to +516 relative to the ATG. These primers were designed to generate a new NcoI site (CCATGG) at the 5' end and a new BclI site (TGATCA) at the 3' end of the amplified open reading frame, respectively. The amplified reaction products were digested with NcoI and BclI, and then separated by electrophoresis on an agarose gel. A 536 base pair fragment corresponding to the desired piece of coreΔ8

was excised and electroeluted from the gel. This fragment was then cloned into the 4601 bp vector fragment of pET-3d (Fig. 78), a vector for bacterial expression of cloned inserts from the T7 promoter (Studier et al., 1990, Meth. Enzymol. 185:60-89). pET system vectors were obtained from Novagen, Inc. (Madison, WI); and cut with NcoI and BamHI (compatible with BclI). The resulting plasmid, pT7/core (Fig. 79), contains the coreΔ8 ORF cloned downstream of the T7 promoter. Restriction mapping was used to determine that the insert in pT7/core had been made so as to retain the NcoI site at the 5' end of the ORF encoding the amino terminus.

Plasmid pT7/core was then transformed into a derivative of *E. coli* strain BL21 (F^- *ompT* $r_B^-{}_{mB}^-$) containing the DE3 lysogen. DE3 is a lambda derivative that carries in the *int* gene a DNA fragment containing the lacI gene, the lacUV5 promoter, the beginning of the lacZ gene, and the gene for T7 RNA polymerase. In a DE3 lysogen, the lacUV5 promoter is inducible by isopropyl-β-D-thiogalactopyranoside (IPTG), such that addition of IPTG to a growing culture of this lysogen induces T7 RNA polymerase (Studier et al., *supra*). The BL21(DE3) lysogen was transformed with pT7/core, and host cells carrying the recombinant pT7/core plasmid were grown in 1 liter of LB broth at 37°C. CoreΔ8 synthesis was induced by the addition of 0.4 mM IPTG, so that the coreΔ8 gene was expressed from the T7 promoter in pT7/core.

This induced culture was harvested, and the coreΔ8 antigen, which is expressed intracellularly, was partially purified. Induced bacterial cells were harvested by low speed centrifugation and resuspended in 20 ml per liter of culture modified lysis and sonication (L&S-1) buffer (50 mM Tris pH 8.0, 10 mM

EDTA, 10 mM dithioerythritol (DTT), 1 mM phenylmethylsulfonyl fluoride (PMSF), 0.25 mg/ml lysozyme, 4 mg of deoxycholate (DOC) per gram of cells). The cell suspension was sonicated three times (Braun-Sonic 2000, half-max. high, 10 seconds) and then centrifuged at 12,000 x *g* for 40 min at 4°C. The supernatant was then treated with 30% ammonium sulfate, stirred for 1 hour at 4°C, and centrifuged at 12,000 x *g* for 40 min. Ammonium sulfate was then added to the supernatant to 40%. This mixture was stirred for 1 hour at 4°C and centrifuged at 12,000 x *g* for 40 min. The precipitated pellet was then resuspended in 6 M urea, 50 mM Tris pH 8.0, 10 mM EDTA, 10 mM DTT, 1 mM PMSF buffer to a protein concentration of 3 mg/ml. This dissolved solution was loaded onto a sizing column (Bio-Rad A1.5, 2.5 x 100 cm), equilibrated and eluted with the same buffer. A protein profile of the eluted fractions was obtained by absorbance at 280 nm, or by a micro protein assay (Bio-Rad). One major peak containing coreΔ8 peptide was observed. Fractions containing this peak were pooled and dialyzed against Tris/saline (10 mM Tris, 0.15 M NaCl), pH 7.5.

The purification steps for the recombinant peptide, expected to be of molecular weight 20.3 kD, were followed by SDS-PAGE analysis on a 15% gel. Samples were mixed with SDS sample buffer to a final concentration of 25 mM Tris-HCl pH 6.8, 4% glycerol, 0.8% SDS, 2% β-mercaptoethanol, 0.0005% bromophenol blue. CoreΔ8 purified in this manner was substituted for the Biodesign core antigen in the immunoassays described in example 8.3 below.

The next step in the construction of pPB-05, the preS1-coreΔ8 fusion vector, was to transfer the coreΔ8 antigen back into a general cloning vector.

pT7/core was digested with XbaI and HindIII to excise the 1056 bp coreΔ8-containing piece, and this piece was subcloned into the 2859 bp vector backbone of pT7T318 which had been digested with XbaI and HindIII so as to reconstruct the XbaI and HindIII sites. The resulting plasmid, pPB-04 (Fig. 80), contains the coreΔ8 antigen surrounded on either end by restriction sites.

To make a fusion of the preS1 region to the 5' end of the coreΔ8 region in pPB-04, the preS1 region was cloned by PCR out of pRO-02. The primers used for this purpose were of sequence 5' ATT CTC GAG GGC ATG GGG ACG 3' (SEQ ID NO:42), which hybridizes from bases -12 to +9 relative to the initiating ATG of preS1, and 5' ACT CCA TGG CCT GAG GAT GA 3' (SEQ ID NO:43), which hybridizes to the opposite strand from bases +331 to +312 relative to the ATG. These primers were designed to generate XhoI (CTCGAG) and NcoI (CCATGG) sites at the amino and carboxyl termini, respectively, of the 343 bp amplified preS1 ORF. The reaction mixture generated by PCR was digested by XhoI, the ends filled in using the Klenow fragment of DNA polymerase I, and the DNA cut by NcoI. The resulting 333 bp fragment containing the preS1 region was gel-purified and cloned into the 3865 bp vector fragment of pPB-04 which had been cut at SmaI in the polylinker region at the 5' end of coreΔ8 and at NcoI at the initiating ATG of the coreΔ8 antigen. Restriction mapping was used to confirm that the inserted DNA was oriented so as to recreate the NcoI site at the fusion point of the preS1 and core regions. The resulting plasmid, pPB-05, contains a fused preS1-coreΔ8 open reading frame surrounded by restriction sites. The nucleotide sequence (SEQ ID NO:44) and predicted amino acid sequence (SEQ ID

NO:45) of the preS1-coreΔ8 fusion region in pPB-05 are shown in Figure 81.

From pPB-05 was made an expression plasmid, pRO-21 (Fig. 82), containing the MS antigen expressed from the p7.5 promoter and the preS1-coreΔ8 antigen expressed from the p11 promoter. pPB-05 was digested with EcoRI and HindIII to excise the preS1-coreΔ8-containing fragment and this fragment was made blunt-ended by Klenow treatment. This 1370 bp insert containing the preS1-coreΔ8 fusion was gel-purified and inserted into the 6072 bp XhoI vector fragment of pRO-11 which had been Klenow-blunted and dephosphorylated by BAP treatment. The orientation of the insert was verified by BamHI digestion to give a restriction map consistent with fragments of 49, 966 and 6427 base pairs, indicating that the preS1-coreΔ8 open reading frame in pRO-21 is oriented so as to be expressible from the p11 promoter.

pPB-09, the plasmid containing the preS1-full-length core fusion, was constructed by first amplifying the full-length core antigen from pCODM/24 by PCR. The primers used for this purpose were of sequence 5' GCG CCA TGG ACA TTG ACC CTT ATA 3' (SEQ ID NO:46), which hybridizes from bases -5 to +19 relative to the initiating ATG of core, and 5' AGT GAA GCT TCC CAC CTT 3' (SEQ ID NO:47), which hybridizes to the opposite strand from bases +592 to +575 relative to the full-length core ATG. These primers were designed to generate NcoI (CCATGG) and HindIII (AAGCTT) sites at the N-terminal and C-terminal ends of core, respectively, upon amplification. The resulting reaction mixture was digested with NcoI and HindIII. The 584 bp core-containing piece was then gel-purified and inserted into the 3181 bp gel-purified vector backbone of pPB-05 from which the

coreΔ8 open reading frame had been removed by NcoI/HindIII digestion. Restriction mapping was used to verify that the insert in the resulting plasmid, pPB-09, was oriented so as to reconstruct both the NcoI and HindIII sites. This process had the effect of switching the full-length core antigen for coreΔ8. The nucleotide sequence (SEQ ID NO:48) and predicted amino acid sequence (SEQ ID NO:49) of the preS1-full-length core fusion in pPB-09 are shown in Figure 83.

An analogous construct to pRO-21 in which the full-length core antigen was substituted for the coreΔ8 antigen was made from pPB-09 and called pRO-23 (Fig. 84). To make pRO-23, the preS1-full-length core insert was purified from pPB-09 by digesting pPB-09 with EcoRI and HindIII, and using Klenow fragment to blunt the ends. This 937 bp fragment was gel-purified and subcloned into the 6072 bp vector fragment of pRO-11 which had been digested with XhoI and the ends filled by Klenow. The correct orientation of the insert in the resulting plasmid, pRO-23, was verified to be such as to allow expression of the fusion from the p11 promoter; digestion with BamHI yielded a restriction map consistent with fragments of 49, 966 and 5994 base pairs. pRO-23, therefore, consists of MS downstream of the p7.5 promoter and the preS1-full-length core antigen fusion downstream of the p11 promoter.

From pPB-05, a plasmid was made for the expression of the preS1-coreΔ8 fusion from the 11 kD promoter. This was done by isolating the 1365 bp EcoRI-HindIII fragment of pPB-05 containing the fusion ORF and blunting the ends with Klenow. This 1370 bp fragment was inserted into the 4149 bp vector fragment of pHTL-5 which had been cut with HindIII and BamHI and made blunt-ended by Klenow. The orientation of

the insert in the resulting plasmid, pHTL-23 (Fig. 85), was determined to be such that the preS1-coreΔ8 should be expressed from the 11 kD promoter; NcoI/StuI digestion yielded a restriction map consistent with fragments of 532 and 4987 bp.

A similar plasmid allowing expression of the preS1-full-length-core antigen from the p11 promoter was constructed from pHTL-23. This was done by cutting pHTL-5 with SacI, treating with mung bean nuclease to blunt the overhanging ends, and then cutting with BglII and isolating the 611 bp fragment of pHTL-5 containing the 3' end encoding the carboxyl terminus of the full-length core antigen. This insert was then ligated into the 4785 bp vector fragment of pHTL-23 which had been digested with NheI, mung bean and BAP-treated, and then digested with BglII to excise the 3' end encoding the carboxyl terminus of the coreΔ8 antigen. The orientation of the insert in the resulting plasmid, pHTL-24 (Fig. 86), was verified to be such that the preS1-full-length-core antigen could be expressed from the p11 promoter; digestion with BglII and StuI yielded a restriction pattern consistent with fragments of 617 and 4779 bp .

Another plasmid, pHTL-18 (Fig. 87), which was made from pRO-23, contains the MS ORF downstream of the modified p7.5 promoter and the preS1-full-length core fusion ORF downstream of the p7.5 promoter. This plasmid was made by isolating a 763 bp fragment containing preS1 fused to the 5' end of the full-length core gene generated by cutting pRO-23 with XhoI, Klenow filling and then digesting with BspEI. This gel-purified insert was then inserted into the 5579 bp vector fragment of pHTL-17 made by cutting with SphI, blunting with mung bean nuclease, and then digesting with BspEI. Orientation of the insert was

determined by restriction mapping to be such that the BspEI site near the 3' end of full-length core had been regenerated. This plasmid contains an insertion of vaccinia DNA between the two promoters and as such is difficult to use to make recombinant viruses, but it is useful for cloning purposes. The region between the two promoters can be deleted to make a plasmid which can be used to generate recombinant viruses.

From pHTL-18 was made an additional plasmid, pHTL-14 (Fig. 88), consisting of the MS antigen downstream of the modified p7.5 promoter, and the preS1-coreΔ8 fusion downstream of the p7.5 promoter. To make this plasmid, the 819 bp core antigen region was isolated from pRO-21 by digestion with NheI, Klenow blunting and NcoI digestion. This fragment was then gel-purified and inserted into the 5290 bp vector fragment of pHTL-18 which had been digested with PvuII and NcoI. The orientation of the insert in pHTL-14 was verified by restriction mapping to be such that the NcoI site at the junction of the preS1 and core ORFs had been regenerated. Like pHTL-17 and pHTL-18, pHTL-14 contains a region of vaccinia DNA between the two promoters. This region should be deleted, and a larger region of the vaccinia TK gene inserted downstream of the preS1-coreΔ8 ORF, in order to use the plasmid for generation of recombinant virus. However, even without these changes, pHTL-14 is useful for cloning purposes.

From pHTL-14, was made an additional plasmid, pHTL-31 (Fig. 89), consisting of MS with the upregulated ATG as described for pHTL-10M above, expressed from the modified p7.5 promoter, and the preS1-coreΔ8 fusion expressed from the p7.5 promoter. To make this plasmid, pHTL-14 was digested with NdeI and BspEI to isolate the region containing p7.5-preS1

and the 5' end of coreΔ8. This 949 bp fragment was gel-purified and inserted into the 5174 bp vector fragment of pHTL-30 which had also been digested with NdeI and BspEI and gel-purified. The insertion in pHTL-31 was made so as to recreate both the NdeI and the BspEI sites.

From pHTL-31 was made an additional vector, pHTL-32 (Fig. 90), consisting of the preS1-coreΔ8 antigen expressed from the p7.5 promoter and the modified p7.5 promoter with no antigen downstream. This plasmid was made by digesting pHTL-31 with XhoI and StuI, Klenow filling, gel-purifying the 5121 bp vector fragment from which the MS gene had been deleted, and self-ligating this plasmid to reclose the circle.

From pHTL-33 and pHTL-31 was also made a plasmid, pHTL-36 (Fig. 75), which is an analog of pHTL-31 in which full-length core has been substituted for coreΔ8. The plasmid pHTL-36, therefore, consists of MS with the upregulated ATG expressed from the modified p7.5 promoter and the fusion preS1-full-length core expressed from the p7.5 promoter. This plasmid was made by digesting pHTL-33 with BspEI and SalI and gel purifying the 149 bp fragment containing the 3' end of the full-length core antigen. This fragment was inserted into the 5995 bp vector fragment of pHTL-31 which had been digested with BspEI and SalI as well. Presence and correct orientation of the insert in pHTL-36 were verified by digestion of the plasmid with BspEI and SalI to yield a restriction map consistent with fragments of sizes 149 and 5995 bp.

7. CONSTRUCTION OF RECOMBINANT TK⁻ VACCINIA VIRUSES

Recombinant vaccinia viruses, for vaccine use, were made by infecting Vero cells (ATCC Accession

No. CCL 81) with wild-type vaccinia virus (New York City Department of Health Laboratories vaccine strain, prepared by Wyeth), and subsequently transfecting the infected cells with a plasmid containing the genes to be inserted flanked by sequences of the vaccinia TK gene, within which homologous recombination occurred. Transfections were performed by one of several methods described in Section 5.3, *supra*. Recombinant viruses were selected by plating on Rat2 cells (ATCC Accession No. CRL 1764) in the presence of BUDR. Each recombinant was purified by three rounds of purification from a single viral plaque, and the identity and purity of each viral stock were verified by commercially available immunoassays for S and e (Abbott Laboratories), and by Southern and western analyses as described in sections below.

8. ANALYSIS OF RECOMBINANT VIRUSES

8.1. SOUTHERN ANALYSIS

A conventional Southern blotting procedure (Southern, 1975, J. Mol. Biol. 98:503) was used to verify the identity and purity of the recombinant viruses which had been plaque-purified. Viral and cellular DNA was isolated using a miniprep procedure from a 100 mm dish of infected host cells. Cells were scraped from the dish into a 15 ml centrifuge tube, spun to pellet, washed in phosphate buffered saline containing 1 mM $MgCl_2$ (PBS-M; PBS = 137 mM NaCl, 3 mM KCl, 8 mM Na_2HPO_4, 1.5 mM KH_2PO_4, pH 7.5), respun and resuspended in 0.6 ml PBS-M. Cells were transferred to a 1.5 ml microcentrifuge tube and the following were added: 30 μl of 10% Triton X-100, 1.5 μl of β-mercaptoethanol, and 48 μl of 250 mM EDTA (pH 8.0). The tubes were vortexed and incubated on ice for 10

minutes with occasional vortexing. Tubes were then spun 800 x g 3 minutes to remove cellular material, the supernatant was transferred to a fresh tube and spun 16,000 x g for 10 minutes to pellet viral cores. The pellet of this second spin was resuspended in 100 μl TE (10 mM Tris, 1 mM EDTA, pH 8.0) and the following were added: 1.5 μl 10 mg/ml proteinase K, 6.7 μl 3 M NaCl, 10 μl 10% SDS, and 0.3 μl β-mercaptoethanol. Tubes were mixed gently and incubated 30 minutes at 55°C. Samples were then extracted twice with an equal volume 1:1 phenol:$CHCl_3$ without vortexing, and ethanol precipitated twice by addition of 4 μl 3 M NaCl and 2.5 volumes ethanol. Precipitated DNA was resuspended in 20 μl H_2O.

An aliquot of this DNA was then digested with a restriction enzyme such as HindIII, size fractionated by electrophoresis on an agarose gel with appropriate markers, and the gel stained and photographed. The DNA in the gel was then transferred to a nylon membrane (Hybond N, Amersham Corp., Arlington Heights, IL) by capillary action. After the transferred DNA was baked onto the membrane, the membrane was prehybridized and hybridized to a labeled nucleotide probe. To probe for the S antigen, a 836 bp fragment of the plasmid pT7T3/S which had been digested with HindIII and EcoRI was gel-purified and nonradioactively labeled by commercially available means (Genius Nonradioactive Labeling and Detection System; Boehringer Mannheim, Indianapolis, IN), in which the DNA was labeled in a random-primed incorporation reaction by the introduction of dUTP which had been conjugated to the steroid hapten digoxigenin. Following hybridization with the labeled probe, the blot was washed and the hybridized probe visualized. Specifically, an alkaline

phosphatase-conjugated antibody to digoxigenin was added and a chemiluminescent alkaline phosphate substrate (Lumiphos, Boehringer Mannheim, Indianapolis, IN) was used to visualize the bound antibody on X-ray film.

DNA from recombinant viruses was analyzed for hybridization to probes for the S, preS2, preS1, and core regions, made by techniques analogous to those described *supra*, from plasmids described *supra* or others; the identity of the recombinants was thus verified. A probe hybridizing within the TK gene of vaccinia virus was also used in assessing the identity and purity of recombinants. Such a probe was made by cutting pGS53 with XbaI and EcoRI to excise the TK region, gel-purifying the 630 bp fragment, and labeling this fragment as described *supra*. Contamination with wild-type vaccinia virus is ruled out by hybridization with such a TK-specific probe, which gives a unique banding pattern for each recombinant or wild-type virus.

8.2. WESTERN BLOT ANALYSIS

The proteins expressed from each recombinant virus were assessed by a modification of the standard western analysis technique (Burnette, 1981, Anal. Biochem. 112:195), using antibodies to the various hepatitis B antigens as probes.

Specifically, to perform a western blot of intracellular proteins probed with antibody to the S region, proteins expressed transiently from recombinant plasmids in infected cells (Rose et al., *supra*) or expressed in cells infected with recombinant viruses were isolated. The procedure for harvesting samples from recombinant-infected cells was as follows. Infected cells were scraped off a 100 mm

dish into a 15 ml centrifuge tube, collected by centrifugation, washed with 1X PBS, resuspended in 1 ml 1X PBS, and disrupted by 3 cycles of freeze/thaw. Samples were then sonicated on an ice/H_2O bath, aliquotted and stored at -70°C. 60 µl of this sample was mixed with 20 µl 4X Laemmli buffer (125 mM Tris pH 6.8, 20% (v/v) glycerol, 4% (w/v) SDS, 10% (v/v) β-mercaptoethanol, 0.0025% (w/v) bromophenol blue), boiled 5 minutes, loaded on a 15% SDS/polyacrylamide gel along with appropriate size standards, and subjected to electrophoresis overnight at 50 V (Laemmli, 1970, Nature 227:680). From each 100 mm dish transient expression experiment, pellets were resuspended in 250 µl 1X PBS and 50 µl of this sample was mixed with 17 µl 4X Laemmli buffer.

The separated proteins were transferred (Protean II apparatus and Trans-blot Cell, BIORAD Laboratories, Richmond, CA) to nitrocellulose (Schleicher & Schuell, Keene, NH) at 70 V for 2 hours in 25 mM Tris pH 8.3, 192 mM glycine, 20% methanol. The membrane containing the transferred proteins was blocked with blocking buffer (2% nonfat dry milk in H_2O) for 1 hour at room temperature and then incubated overnight at room temperature in 25 ml blocking buffer containing 1/250 rabbit anti-HBsAg polyclonal antibody (cat. no. KM63P, Accurate Chemical & Scientific Corp., Westbury, NY). The blot was then washed 3 X 10 minutes in 0.1% Tween 20, 100 mM Tris pH 7.5, 1.5 M NaCl, 0.02% NaN_3. Next, a goat anti-rabbit IgG alkaline phosphatase-conjugated second antibody (cat. no. 8612-0081, Cappel Research Products, Durham, NC) was diluted 1/10,000 in blocking buffer and incubated with the blot for 2 hours at room temperature. The blot was then washed 3 X 10 minutes as above. Color was developed by addition of alkaline phosphatase

substrate (NBT/BCIP, Promega Protoblot System, Promega, Madison, WI) in alkaline phosphatase buffer (100 mM Tris pH 9.5, 100 mM NaCl, 5 mM $MgCl_2$).

A representative Western blot of polypeptides in the cell pellet of a number of recombinant viruses probed with anti-S antibodies is shown in Figure 91. It is apparent from this analysis and other similar gels that in some constructs designed to express the MS antigen, significant expression of a polypeptide of the correct size for S occurs as well, and that the relative distribution of MS and S varies among the constructs. In the constructs represented in this western blot, no care was taken to match the sequence upstream of the initiating ATG of MS to the Kozak consensus sequence for ribosome binding described *supra*. Analysis of constructs such as pHTL-10M in which the consensus sequence had been matched did indicate a higher MS to S ratio. The smaller surface proteins, MS and S, were also seen in some of the LS-producing viruses, particularly pGS52/S1. A similar approach can be taken in the LS constructs to try to upregulate the LS initiation codon.

With respect to two independent recombinant vaccinia viruses derived from pHTL-28 (encoding core-S* operably linked to modified p7.5) and purified, the production of e sequences was detected by enzyme-linked immunoassay (ELISA) procedures using anti-e antibody, but no protein products reactive with anti-S antibody were detectable by ELISA. Southern analysis of the viral DNA confirmed the presence of both core and S region sequences. We also performed western analyses on the proteins isolated from cells infected with the purified recombinants. Both the anti-core and anti-S antibodies appeared to detect

weakly the same two bands in each blot, the larger of which was about the expected size for the core-S* fusion protein. This suggests that these recombinant viruses do make some form of fusion, but that perhaps the epitopes are in an unusual conformation and are harder to detect. We conclude that perhaps when we make a fusion to core of the piece of S contained in pHTL-18 (which should include the dominant a epitope), the conformation of S is changed such that it is no longer detectable, or at least not easily detectable, with the standard anti-S antibodies.

8.3. IMMUNOASSAYS FOR DETECTION OF ANTIBODIES TO THE HBV ANTIGENS AND TO VACCINIA VIRUS

Enzyme linked immunoassay (ELISA) procedures were used to detect antibodies directed to vaccinia virus coat proteins and to HBV S, preS1, preS2 and core regions in the serum of mice which had been immunized with recombinant viruses. The amount of HBV-specific antibody present in test serum was quantified by interpolation from a standard curve established using serial dilutions of a monoclonal antibody of known titer specific for HBV S, preS1, preS2, or core regions.

Microtiter plates were coated with purified wild-type vaccinia virus, core protein, plasma-derived HBsAg, a synthetic peptide derived from preS1, or a synthetic peptide derived from preS2. After coating with antigen, the microtiter plates were rinsed with water and then blocked with non-fat milk. Plates were then washed with a Tris buffer (10 mM Tris pH 7.5, 0.15 M NaCl, 0.5% Tween-20).

50 μl of serum sample or antibody standard were then added to triplicate antigen-coated, blocked wells. The serum dilution used was 1:100 in Tris/saline (10 mM Tris pH 7.5, 150 mM NaCl) plus 1%

fetal bovine serum. The reaction was allowed to proceed for 90 minutes at 37°C, and the plates were washed to remove unbound antibody. Antigen-bound antibody was detected by the addition of polyclonal goat anti-mouse antibody conjugated with alkaline phosphatase. After incubation for 60 minutes at 37°C, the plates were washed to remove unbound conjugate and the color reaction was developed. After 15 to 30 minutes, dilute sodium hydroxide was added to stop the reaction and the absorbance at 405 nm was read in an ELISA photometer. The best-fit polynomial equation was calculated for the standards for each assay, and this curve was then used to solve for the concentration of antibody associated with each OD_{405} value for the mouse serum samples.

8.4. PRELIMINARY DOSE AND STRAIN ANALYSES

Mice are often used in initial attempts to assess the immunogenicity of potential vaccines. The immune response of mice to the hepatitis B surface antigens is known to vary depending on the H-2 haplotype of the strain (Milich et al., 1984, J. Exp. Med. 159:41; Milich et al., 1985, Proc. Natl. Acad. Sci. USA 82:8168). However, we were unable to find any published data on the effect of haplotype on the response of mice to hepatitis B antigens when expressed from vaccinia virus. We therefore performed an experiment to assess the importance of the mouse strain and dose used. Seven week old female mice of either C57BL/6 or BALB/c strain were inoculated intraperitoneally with 0.5 ml of a 1:1 mixture of recombinant viruses made from pGS53/S2 and pRO-18. The combination of these two viruses should result in the expression of all of the following epitopes: core, preS1, preS2, and S; other combinations of

viruses could also have been used. Total doses of recombinant vaccinia virus of 1 X 10^7 or 1 X 10^8 plaque forming units (pfu, as determined by titration on Vero cells) were tested in two separate groups (all groups had 10 mice). Two groups of control animals received equivalent doses of wild-type virus. Mice were bled weekly for 16 weeks and pooled sera from each group were tested for antibodies to the S and core antigens using the immunoassays described in example Section 8.3, above. Anti-vaccinia responses were seen in all vaccinia-inoculated mice. Strong anti-S responses were seen in both groups of mice which received the mixture of recombinants at 1 X 10^8 pfu, with the response in the BALB/c mice comparable to that generated in response to 1 μg of the Merck Recombivax vaccine, and the anti-S response in the C57BL/6 slightly lower in magnitude. Anti-core responses were seen in both groups of mice which had been inoculated with the recombinant viruses at 1 X 10^8, as well. Anti-preS2 and weak anti-preS1 responses were seen in the BALB/c mice at 1 X 10^8 pfu of recombinant virus mixture, as well. Based on this experiment, BALB/c mice were chosen as the better strain in which to pursue further immunogenicity studies. It is also desirable to test the ability of the viruses described *supra* to overcome nonresponsiveness to S in mouse strains which are genetic S-nonresponders (for example the H-2^s haplotype). Since no suitable congenic strain to BALB/c with an H-2^s haplotype was readily available, an additional experiment testing the dose response of A.BY mice (H-2^b) was undertaken. These mice, for which a suitable H-2^s congenic is available (A.SW) were tested for tolerance of vaccinia virus and for antibody response to the S-expressing recombinant virus made from pGS53/S at doses of 5 x 10^8 and 1 x 10^9

pfu. A good anti-S response was observed at both doses, so that this strain can be used to test the ability of recombinant viruses to overcome S-nonresponsiveness in mice.

8.5. IMMUNOGENICITY ASSESSMENT IN MICE

Once a mouse strain and dose of virus have been chosen, the immunogenicity of a number of individual viruses and combinations of viruses are compared in mice. The goal is to discover the combination of antigens which gives the best immune response to all the regions (core, preS1, preS2, and S). The viruses/combinations of viruses that are tested are as follows (other combinations can also be used; the ratio of one virus to the other can be varied to change the ratio of antigens expressed in each candidate vaccine). Doses of vaccinia virus in the total range of 10^8 to 10^9 total pfu are tested. Group 1 receives a mixture of viruses (1:1 or other) made from pRO-18 (containing promoter-antigen combinations p7.5-LS & p11-coreΔ8) and pGS53/S2 (p7.5-MS). Group 2 receives a virus made from pHTL-30 (p7.5-coreΔ8 & modified p7.5-MS), and group 3 receives a virus made from pHTL-31 (p7.5-preS1-coreΔ8 and modified p7.5-MS). Group 4 receives a mixture of viruses made from pHTL-26 (modified p7.5-core-preS1*), pHTL-27 (modified p7.5-core-preS2), and pHTL-8 (modified p7.5-S). Group 5 receives a mixture of viruses made from pHTL-26 and pHTL-10M (modified p7.5-MS). Group 6 receives a mixture of viruses made from pHTL-27 and pHTL-8. Group 7 receives a virus made from pHTL-35 (modified p7.5-LS and p7.5-full-length core). Control groups receive viruses from plasmids pHTL-32 (p7.5-preS1-coreΔ8), pHTL-10M, pHTL-8, and pSC10 (p11-β-galactosidase), respectively.

Additional control groups separately receive peptides, such as those described for preS1 and preS2 in Example 8.3, above, (or in the case of S, the S-only vaccine, Merck's Recombivax, the whole S protein) corresponding to the immunodominant epitopes of S, preS2, preS1, and core in the presence of adjuvant such as alum.

As in the earlier dose and strain experiments, the immune response of the mice to these various combinations of antigens are assessed by use of the immunoassays described above. One or several viruses or combinations of viruses are chosen for further study based on the strength of the anti-HBV immune response which they generate. Further analysis can include testing in strains of mice which are known to lack the ability to produce an antibody response to S antigen by itself ($H\text{-}2^s$ mice), to see if the addition of other HBV epitopes such as preS1, preS2 and core can overcome S-nonresponsiveness in these mice. These same viruses can be analyzed in chimpanzees for the ability to protect the animals from challenge with hepatitis B virus. A vaccine which proves efficacious in chimpanzees can be tested in human clinical trials.

9. PRELIMINARY RESULTS USING INSERTION SITES IN VACCINIA NON-CODING REGIONS

Three plasmid insertion vectors containing HBV sequences flanked by vaccinia non-coding regions were constructed, for targeting insertion into the F14L-F15L, C12L-C11R, and A53R-A55R regions of the vaccinia genome, respectively. When two of the constructs (regions F14L-F15L and C12L-C11R, respectively) were used in attempts to generate recombinant vaccinia viruses, no recombinants were obtained out of 10,000 viruses screened (using the blue-white screening method based on β-galactoside expression). These recombinants would have been

identifiable because they form blue plaques when overlaid with agarose containing X-gal. With the third construct (A53R-A55R), three blue recombinants were found out of 10,000 screened, but when we attempted to purify these viruses by picking the blue plaques, growing up a small stock, plating again and overlaying again, we found that no blue plaques remained. We tentatively conclude that the recombinant was either unstable or at a severe growth disadvantage relative to the non-recombinant virus. However, the strategy we used to select these sites can be extended to other noncoding areas. The frequency of recombinants might be increased by increasing the length of the areas of flanking vaccinia DNA in the plasmid, perhaps up to about 500 bp on each side. Use of such large flanking regions may necessitate inclusion of some DNA in these flanking regions from the nonessential genes surrounding the targeted insertion site.

10. DEPOSIT OF MICROORGANISMS

Bacteria carrying plasmid pHTL-31 encoding HBV preS1-core(Δ8) operably linked to the p7.5 promoter, and HBV MS (with an upregulated ATG) operably linked to the modified p7.5 promoter were deposited on November 12, 1992 with the American Type Culture Collection (ATCC), 1201 Parklawn Drive, Rockville, Maryland 20852, under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedures, and assigned accession no. 69124.

The present invention is not to be limited in scope by the microorganisms deposited or the specific embodiments described herein. Indeed,

various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description and accompanying figures. Such modifications are intended to fall within the scope of the appended claims.

Various publications are cited herein, the disclosures of which are incorporated by reference in their entireties.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Souw, Peter T.S.
O'Keefe, Rhonda W.
Lewis, Tatyana
Bernstine, Edward G.

(ii) TITLE OF INVENTION: Hepatitis B Virus Vaccines

(iii) NUMBER OF SEQUENCES: 59

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Pennie & Edmonds
(B) STREET: 1155 Avenue of the Americas
(C) CITY: New York,
(D) STATE: New York
(E) COUNTRY: U.S.A.
(F) ZIP: 10036-2711

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: To be assigned
(B) FILING DATE: On even date herewith
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Misrock, S. Leslie
(B) REGISTRATION NUMBER: 18,872
(C) REFERENCE/DOCKET NUMBER: 7441-003

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (212)790-9090
(B) TELEFAX: (212)869-8864-9741
(C) TELEX: 66141 PENNIE

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(C) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Pro Arg Arg Arg Arg Ser Gln Ser
1 5

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 64 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ATAAATAATA AATACAATAA TTAATTTCTC GTAAAAGTAG AAAATATATT CTAATTTATT 60

GCAC 64

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 104 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CTAGAAGCGA TGCTACGCTA GTCACAATCA CCACTTTCAT ATTTAGAATA TATGTATGTA 60

AAAATATAGT AGAATTTCAT TTTGTTTTTT TCTATGCTAT AAAT 104

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

AAAAATTGAA ATTTTATTTT TTTTTTTTGG AATATAAATA AG 42

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1170 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..1170

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
ATG GGG ACG AAT CTT TCT GTT CCC AAT CCT CTG GGA TTC TTT CCC GAT      48
Met Gly Thr Asn Leu Ser Val Pro Asn Pro Leu Gly Phe Phe Pro Asp
 1               5                  10                  15

CAT CAG TTG GAC CCT GCA TTC GGA GCC AAC TCA AAC AAT CCA GAT TGG      96
His Gln Leu Asp Pro Ala Phe Gly Ala Asn Ser Asn Asn Pro Asp Trp
            20                  25                  30

GAC TTC AAC CCC ATC AAG GAC CAC TGG CCA GCA GCC AAC CAG GTA GGA     144
Asp Phe Asn Pro Ile Lys Asp His Trp Pro Ala Ala Asn Gln Val Gly
        35                  40                  45
```

```
GTG GGA GCA TTC GGG CCA GGG TTC ACC CCT CCA CAC GGC GGT GTT TTG          192
Val Gly Ala Phe Gly Pro Gly Phe Thr Pro Pro His Gly Gly Val Leu
     50                  55                  60

GGG TGG AGC CCT CAG GCT CAG GGC ATA TTG ACC ACA GTG TCA ACA ATT          240
Gly Trp Ser Pro Gln Ala Gln Gly Ile Leu Thr Thr Val Ser Thr Ile
 65                  70                  75                  80

CCT CCT CCT GCC TCC ACC AAT CGG CAG TCA GGA AGG CAG CCT ACT CCC          288
Pro Pro Pro Ala Ser Thr Asn Arg Gln Ser Gly Arg Gln Pro Thr Pro
                 85                  90                  95

ATC TCT CCA CCT CTA AGA GAC AGT CAT CCT CAG GCC ATG CAG TGG AAC          336
Ile Ser Pro Pro Leu Arg Asp Ser His Pro Gln Ala Met Gln Trp Asn
            100                 105                 110

TCC ACT GCC TTC CAC CAA GCT CTG CAG GAT CCC AGA GTC AGG GGT CTG          384
Ser Thr Ala Phe His Gln Ala Leu Gln Asp Pro Arg Val Arg Gly Leu
        115                 120                 125

TAT TTT CCT GCT GGT GGC TCC AGT TCA GGA ACA GTA AAC CCT GCT CCG          432
Tyr Phe Pro Ala Gly Gly Ser Ser Ser Gly Thr Val Asn Pro Ala Pro
    130                 135                 140

AAT ATT GCC TCT CAC ATC TCG TCA ATC TCC GCG AGG ACT GGG GAC CCT          480
Asn Ile Ala Ser His Ile Ser Ser Ile Ser Ala Arg Thr Gly Asp Pro
145                 150                 155                 160

GTG ACG AAC ATG GAG AAC ATC ACA TCA GGA TTC CTA GGA CCC CTG CTC          528
Val Thr Asn Met Glu Asn Ile Thr Ser Gly Phe Leu Gly Pro Leu Leu
                165                 170                 175

GTG TTA CAG GCG GGG TTT TTC TTG TTG ACA AGA ATC CTC ACA ATA CCG          576
Val Leu Gln Ala Gly Phe Phe Leu Leu Thr Arg Ile Leu Thr Ile Pro
            180                 185                 190

CAG AGT CTA GAC TCG TGG TGG ACT TCT CTC AAT TTT CTA GGG GGA TCA          624
Gln Ser Leu Asp Ser Trp Trp Thr Ser Leu Asn Phe Leu Gly Gly Ser
        195                 200                 205

CCC GTG TGT CTT GGC CAA AAT TCG CAG TCC CCA ACC TCC AAT CAC TCA          672
Pro Val Cys Leu Gly Gln Asn Ser Gln Ser Pro Thr Ser Asn His Ser
    210                 215                 220

CCA ACC TCC TGT CCT CCA ATT TGT CCT GGT TAT CGC TGG ATG TGT CTG          720
Pro Thr Ser Cys Pro Pro Ile Cys Pro Gly Tyr Arg Trp Met Cys Leu
225                 230                 235                 240

CGG CGT TTT ATC ATA TTC CTC TTC ATC CTG CTG CTA TGC CTC ATC TTC          768
Arg Arg Phe Ile Ile Phe Leu Phe Ile Leu Leu Leu Cys Leu Ile Phe
                245                 250                 255

TTA TTG GTT CTT CTG GAT TAT CAA GGT ATG TTG CCC GTT TGT CCT CTA          816
Leu Leu Val Leu Leu Asp Tyr Gln Gly Met Leu Pro Val Cys Pro Leu
            260                 265                 270

ATT CCA GGA TCA ACA ACA ACC AGT ACG GGA CCA TGC AAA ACC TGC ACG          864
Ile Pro Gly Ser Thr Thr Thr Ser Thr Gly Pro Cys Lys Thr Cys Thr
        275                 280                 285

ACT CCT GCT CAA GGC AAC TCT ATG TTT CCC TCA TGT TGC TGT ACA AAA          912
Thr Pro Ala Gln Gly Asn Ser Met Phe Pro Ser Cys Cys Cys Thr Lys
    290                 295                 300

CCT ACG GAT GGA AAT TGC ACC TGT ATT CCC ATC CCA TCG TCT TGG GCT          960
Pro Thr Asp Gly Asn Cys Thr Cys Ile Pro Ile Pro Ser Ser Trp Ala
305                 310                 315                 320

TTC GCA AAA TAC CTA TGG GAG TGG GCC TCA GTC CGT TTC TCT TGG CTC         1008
Phe Ala Lys Tyr Leu Trp Glu Trp Ala Ser Val Arg Phe Ser Trp Leu
```

```
                  325                 330                 335

AGT TTA CTA GTG CCA TTT GTT CAG TGG TTC GTA GGG CTT TCC CCC ACT      1056
Ser Leu Leu Val Pro Phe Val Gln Trp Phe Val Gly Leu Ser Pro Thr
            340                 345                 350

GTT TGG CTT TCA GCT ATA TGG ATG ATG TGG TAT TGG GGG CCA AGT CTG      1104
Val Trp Leu Ser Ala Ile Trp Met Met Trp Tyr Trp Gly Pro Ser Leu
        355                 360                 365

TAC AGC ATC GTG AGT CCC TTT ATA CCG CTG TTA CCA ATT TTC TTT TGT      1152
Tyr Ser Ile Val Ser Pro Phe Ile Pro Leu Leu Pro Ile Phe Phe Cys
    370                 375                 380

CTC TGG GTA TAC ATT TAA                                             1170
Leu Trp Val Tyr Ile
385                 390
```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 389 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
Met Gly Thr Asn Leu Ser Val Pro Asn Pro Leu Gly Phe Phe Pro Asp
  1               5                  10                  15

His Gln Leu Asp Pro Ala Phe Gly Ala Asn Ser Asn Asn Pro Asp Trp
            20                  25                  30

Asp Phe Asn Pro Ile Lys Asp His Trp Pro Ala Ala Asn Gln Val Gly
        35                  40                  45

Val Gly Ala Phe Gly Pro Gly Phe Thr Pro Pro His Gly Gly Val Leu
    50                  55                  60

Gly Trp Ser Pro Gln Ala Gln Gly Ile Leu Thr Thr Val Ser Thr Ile
65                  70                  75                  80

Pro Pro Pro Ala Ser Thr Asn Arg Gln Ser Gly Arg Gln Pro Thr Pro
                85                  90                  95

Ile Ser Pro Pro Leu Arg Asp Ser His Pro Gln Ala Met Gln Trp Asn
            100                 105                 110

Ser Thr Ala Phe His Gln Ala Leu Gln Asp Pro Arg Val Arg Gly Leu
        115                 120                 125

Tyr Phe Pro Ala Gly Gly Ser Ser Ser Gly Thr Val Asn Pro Ala Pro
    130                 135                 140

Asn Ile Ala Ser His Ile Ser Ser Ile Ser Ala Arg Thr Gly Asp Pro
145                 150                 155                 160

Val Thr Asn Met Glu Asn Ile Thr Ser Gly Phe Leu Gly Pro Leu Leu
                165                 170                 175

Val Leu Gln Ala Gly Phe Phe Leu Leu Thr Arg Ile Leu Thr Ile Pro
            180                 185                 190

Gln Ser Leu Asp Ser Trp Trp Thr Ser Leu Asn Phe Leu Gly Gly Ser
        195                 200                 205

Pro Val Cys Leu Gly Gln Asn Ser Gln Ser Pro Thr Ser Asn His Ser
```

```
     210                 215                 220

Pro Thr Ser Cys Pro Pro Ile Cys Pro Gly Tyr Arg Trp Met Cys Leu
225                 230                 235                 240

Arg Arg Phe Ile Ile Phe Leu Phe Ile Leu Leu Leu Cys Leu Ile Phe
                245                 250                 255

Leu Leu Val Leu Leu Asp Tyr Gln Gly Met Leu Pro Val Cys Pro Leu
            260                 265                 270

Ile Pro Gly Ser Thr Thr Thr Ser Thr Gly Pro Cys Lys Thr Cys Thr
        275                 280                 285

Thr Pro Ala Gln Gly Asn Ser Met Phe Pro Ser Cys Cys Cys Thr Lys
    290                 295                 300

Pro Thr Asp Gly Asn Cys Thr Cys Ile Pro Ile Pro Ser Ser Trp Ala
305                 310                 315                 320

Phe Ala Lys Tyr Leu Trp Glu Trp Ala Ser Val Arg Phe Ser Trp Leu
                325                 330                 335

Ser Leu Leu Val Pro Phe Val Gln Trp Phe Val Gly Leu Ser Pro Thr
            340                 345                 350

Val Trp Leu Ser Ala Ile Trp Met Met Trp Tyr Trp Gly Pro Ser Leu
        355                 360                 365

Tyr Ser Ile Val Ser Pro Phe Ile Pro Leu Leu Pro Ile Phe Phe Cys
    370                 375                 380

Leu Trp Val Tyr Ile
385
```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 846 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..846

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```
ATG CAG TGG AAC TCC ACT GCC TTC CAC CAA GCT CTG CAG GAT CCC AGA      48
Met Gln Trp Asn Ser Thr Ala Phe His Gln Ala Leu Gln Asp Pro Arg
  1               5                  10                  15

GTC AGG GGT CTG TAT TTT CCT GCT GGT GGC TCC AGT TCA GGA ACA GTA      96
Val Arg Gly Leu Tyr Phe Pro Ala Gly Gly Ser Ser Ser Gly Thr Val
            20                  25                  30

AAC CCT GCT CCG AAT ATT GCC TCT CAC ATC TCG TCA ATC TCC GCG AGG     144
Asn Pro Ala Pro Asn Ile Ala Ser His Ile Ser Ser Ile Ser Ala Arg
        35                  40                  45

ACT GGG GAC CCT GTG ACG AAC ATG GAG AAC ATC ACA TCA GGA TTC CTA     192
Thr Gly Asp Pro Val Thr Asn Met Glu Asn Ile Thr Ser Gly Phe Leu
    50                  55                  60
```

```
GGA CCC CTG CTC GTG TTA CAG GCG GGG TTT TTC TTG TTG ACA AGA ATC        240
Gly Pro Leu Leu Val Leu Gln Ala Gly Phe Phe Leu Leu Thr Arg Ile
 65                  70                  75                  80

CTC ACA ATA CCG CAG AGT CTA GAC TCG TGG TGG ACT TCT CTC AAT TTT        288
Leu Thr Ile Pro Gln Ser Leu Asp Ser Trp Trp Thr Ser Leu Asn Phe
                85                  90                  95

CTA GGG GGA TCA CCC GTG TGT CTT GGC CAA AAT TCG CAG TCC CCA ACC        336
Leu Gly Gly Ser Pro Val Cys Leu Gly Gln Asn Ser Gln Ser Pro Thr
            100                 105                 110

TCC AAT CAC TCA CCA ACC TCC TGT CCT CCA ATT TGT CCT GGT TAT CGC        384
Ser Asn His Ser Pro Thr Ser Cys Pro Pro Ile Cys Pro Gly Tyr Arg
        115                 120                 125

TGG ATG TGT CTG CGG CGT TTT ATC ATA TTC CTC TTC ATC CTG CTG CTA        432
Trp Met Cys Leu Arg Arg Phe Ile Ile Phe Leu Phe Ile Leu Leu Leu
    130                 135                 140

TGC CTC ATC TTC TTA TTG GTT CTT CTG GAT TAT CAA GGT ATG TTG CCC        480
Cys Leu Ile Phe Leu Leu Val Leu Leu Asp Tyr Gln Gly Met Leu Pro
145                 150                 155                 160

GTT TGT CCT CTA ATT CCA GGA TCA ACA ACA ACC AGT ACG GGA CCA TGC        528
Val Cys Pro Leu Ile Pro Gly Ser Thr Thr Thr Ser Thr Gly Pro Cys
                165                 170                 175

AAA ACC TGC ACG ACT CCT GCT CAA GGC AAC TCT ATG TTT CCC TCA TGT        576
Lys Thr Cys Thr Thr Pro Ala Gln Gly Asn Ser Met Phe Pro Ser Cys
            180                 185                 190

TGC TGT ACA AAA CCT ACG GAT GGA AAT TGC ACC TGT ATT CCC ATC CCA        624
Cys Cys Thr Lys Pro Thr Asp Gly Asn Cys Thr Cys Ile Pro Ile Pro
        195                 200                 205

TCG TCT TGG GCT TTC GCA AAA TAC CTA TGG GAG TGG GCC TCA GTC CGT        672
Ser Ser Trp Ala Phe Ala Lys Tyr Leu Trp Glu Trp Ala Ser Val Arg
    210                 215                 220

TTC TCT TGG CTC AGT TTA CTA GTG CCA TTT GTT CAG TGG TTC GTA GGG        720
Phe Ser Trp Leu Ser Leu Leu Val Pro Phe Val Gln Trp Phe Val Gly
225                 230                 235                 240

CTT TCC CCC ACT GTT TGG CTT TCA GCT ATA TGG ATG ATG TGG TAT TGG        768
Leu Ser Pro Thr Val Trp Leu Ser Ala Ile Trp Met Met Trp Tyr Trp
                245                 250                 255

GGG CCA AGT CTG TAC AGC ATC GTG AGT CCC TTT ATA CCG CTG TTA CCA        816
Gly Pro Ser Leu Tyr Ser Ile Val Ser Pro Phe Ile Pro Leu Leu Pro
            260                 265                 270

ATT TTC TTT TGT CTC TGG GTA TAC ATT TAA                               846
Ile Phe Phe Cys Leu Trp Val Tyr Ile
        275                 280
```

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 281 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```
Met Gln Trp Asn Ser Thr Ala Phe His Gln Ala Leu Gln Asp Pro Arg
 1               5                  10                  15
```

```
Val Arg Gly Leu Tyr Phe Pro Ala Gly Gly Ser Ser Ser Gly Thr Val
            20                  25                  30

Asn Pro Ala Pro Asn Ile Ala Ser His Ile Ser Ser Ile Ser Ala Arg
        35                  40                  45

Thr Gly Asp Pro Val Thr Asn Met Glu Asn Ile Thr Ser Gly Phe Leu
    50                  55                  60

Gly Pro Leu Leu Val Leu Gln Ala Gly Phe Phe Leu Leu Thr Arg Ile
65                  70                  75                  80

Leu Thr Ile Pro Gln Ser Leu Asp Ser Trp Trp Thr Ser Leu Asn Phe
                85                  90                  95

Leu Gly Gly Ser Pro Val Cys Leu Gly Gln Asn Ser Gln Ser Pro Thr
            100                 105                 110

Ser Asn His Ser Pro Thr Ser Cys Pro Pro Ile Cys Pro Gly Tyr Arg
        115                 120                 125

Trp Met Cys Leu Arg Arg Phe Ile Ile Phe Leu Phe Ile Leu Leu Leu
    130                 135                 140

Cys Leu Ile Phe Leu Leu Val Leu Leu Asp Tyr Gln Gly Met Leu Pro
145                 150                 155                 160

Val Cys Pro Leu Ile Pro Gly Ser Thr Thr Thr Ser Thr Gly Pro Cys
                165                 170                 175

Lys Thr Cys Thr Thr Pro Ala Gln Gly Asn Ser Met Phe Pro Ser Cys
            180                 185                 190

Cys Cys Thr Lys Pro Thr Asp Gly Asn Cys Thr Cys Ile Pro Ile Pro
        195                 200                 205

Ser Ser Trp Ala Phe Ala Lys Tyr Leu Trp Glu Trp Ala Ser Val Arg
    210                 215                 220

Phe Ser Trp Leu Ser Leu Leu Val Pro Phe Val Gln Trp Phe Val Gly
225                 230                 235                 240

Leu Ser Pro Thr Val Trp Leu Ser Ala Ile Trp Met Met Trp Tyr Trp
                245                 250                 255

Gly Pro Ser Leu Tyr Ser Ile Val Ser Pro Phe Ile Pro Leu Leu Pro
            260                 265                 270

Ile Phe Phe Cys Leu Trp Val Tyr Ile
        275                 280
```

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 681 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..681

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```
ATG GAG AAC ATC ACA TCA GGA TTC CTA GGA CCC CTG CTC GTG TTA CAG          48
Met Glu Asn Ile Thr Ser Gly Phe Leu Gly Pro Leu Leu Val Leu Gln
  1               5                  10                  15

GCG GGG TTT TTC TTG TTG ACA AGA ATC CTC ACA ATA CCG CAG AGT CTA          96
Ala Gly Phe Phe Leu Leu Thr Arg Ile Leu Thr Ile Pro Gln Ser Leu
            20                  25                  30

GAC TCG TGG TGG ACT TCT CTC AAT TTT CTA GGG GGA TCA CCC GTG TGT         144
Asp Ser Trp Trp Thr Ser Leu Asn Phe Leu Gly Gly Ser Pro Val Cys
        35                  40                  45

CTT GGC CAA AAT TCG CAG TCC CCA ACC TCC AAT CAC TCA CCA ACC TCC         192
Leu Gly Gln Asn Ser Gln Ser Pro Thr Ser Asn His Ser Pro Thr Ser
    50                  55                  60

TGT CCT CCA ATT TGT CCT GGT TAT CGC TGG ATG TGT CTG CGG CGT TTT         240
Cys Pro Pro Ile Cys Pro Gly Tyr Arg Trp Met Cys Leu Arg Arg Phe
65                  70                  75                  80

ATC ATA TTC CTC TTC ATC CTG CTG CTA TGC CTC ATC TTC TTA TTG GTT         288
Ile Ile Phe Leu Phe Ile Leu Leu Leu Cys Leu Ile Phe Leu Leu Val
                85                  90                  95

CTT CTG GAT TAT CAA GGT ATG TTG CCC GTT TGT CCT CTA ATT CCA GGA         336
Leu Leu Asp Tyr Gln Gly Met Leu Pro Val Cys Pro Leu Ile Pro Gly
            100                 105                 110

TCA ACA ACA ACC AGT ACG GGA CCA TGC AAA ACC TGC ACG ACT CCT GCT         384
Ser Thr Thr Thr Ser Thr Gly Pro Cys Lys Thr Cys Thr Thr Pro Ala
        115                 120                 125

CAA GGC AAC TCT ATG TTT CCC TCA TGT TGC TGT ACA AAA CCT ACG GAT         432
Gln Gly Asn Ser Met Phe Pro Ser Cys Cys Cys Thr Lys Pro Thr Asp
    130                 135                 140

GGA AAT TGC ACC TGT ATT CCC ATC CCA TCG TCT TGG GCT TTC GCA AAA         480
Gly Asn Cys Thr Cys Ile Pro Ile Pro Ser Ser Trp Ala Phe Ala Lys
145                 150                 155                 160

TAC CTA TGG GAG TGG GCC TCA GTC CGT TTC TCT TGG CTC AGT TTA CTA         528
Tyr Leu Trp Glu Trp Ala Ser Val Arg Phe Ser Trp Leu Ser Leu Leu
                165                 170                 175

GTG CCA TTT GTT CAG TGG TTC GTA GGG CTT TCC CCC ACT GTT TGG CTT         576
Val Pro Phe Val Gln Trp Phe Val Gly Leu Ser Pro Thr Val Trp Leu
            180                 185                 190

TCA GCT ATA TGG ATG ATG TGG TAT TGG GGG CCA AGT CTG TAC AGC ATC         624
Ser Ala Ile Trp Met Met Trp Tyr Trp Gly Pro Ser Leu Tyr Ser Ile
        195                 200                 205

GTG AGT CCC TTT ATA CCG CTG TTA CCA ATT TTC TTT TGT CTC TGG GTA         672
Val Ser Pro Phe Ile Pro Leu Leu Pro Ile Phe Phe Cys Leu Trp Val
    210                 215                 220

TAC ATT TAA                                                             681
Tyr Ile
225
```

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 226 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```
Met Glu Asn Ile Thr Ser Gly Phe Leu Gly Pro Leu Leu Val Leu Gln
 1               5                  10                  15

Ala Gly Phe Phe Leu Leu Thr Arg Ile Leu Thr Ile Pro Gln Ser Leu
            20                  25                  30

Asp Ser Trp Trp Thr Ser Leu Asn Phe Leu Gly Gly Ser Pro Val Cys
        35                  40                  45

Leu Gly Gln Asn Ser Gln Ser Pro Thr Ser Asn His Ser Pro Thr Ser
    50                  55                  60

Cys Pro Pro Ile Cys Pro Gly Tyr Arg Trp Met Cys Leu Arg Arg Phe
65                  70                  75                  80

Ile Ile Phe Leu Phe Ile Leu Leu Leu Cys Leu Ile Phe Leu Leu Val
                85                  90                  95

Leu Leu Asp Tyr Gln Gly Met Leu Pro Val Cys Pro Leu Ile Pro Gly
            100                 105                 110

Ser Thr Thr Thr Ser Thr Gly Pro Cys Lys Thr Cys Thr Thr Pro Ala
        115                 120                 125

Gln Gly Asn Ser Met Phe Pro Ser Cys Cys Cys Thr Lys Pro Thr Asp
    130                 135                 140

Gly Asn Cys Thr Cys Ile Pro Ile Pro Ser Ser Trp Ala Phe Ala Lys
145                 150                 155                 160

Tyr Leu Trp Glu Trp Ala Ser Val Arg Phe Ser Trp Leu Ser Leu Leu
                165                 170                 175

Val Pro Phe Val Gln Trp Phe Val Gly Leu Ser Pro Thr Val Trp Leu
            180                 185                 190

Ser Ala Ile Trp Met Met Trp Tyr Trp Gly Pro Ser Leu Tyr Ser Ile
        195                 200                 205

Val Ser Pro Phe Ile Pro Leu Leu Pro Ile Phe Phe Cys Leu Trp Val
    210                 215                 220

Tyr Ile
225
```

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 nucleotides
(B) TYPE: nucleic acid
(C) SRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGAGGATCCA GCATCAAGG 19

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 534 nucleotides
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..534

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
ATG GAC ATT GAC CCT TAT AAA GAA TTT GGA GCT ACT GTG GAG TTA CTC      48
Met Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ala Thr Val Glu Leu Leu
 1               5                  10                  15

TCG TTT TTG CCT TCT GAC TTC TTT CCT TCC GTC AGA GAT CTC CTA GAC      96
Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser Val Arg Asp Leu Leu Asp
            20                  25                  30

ACC GCC TCA GCT CTG TAT CGG GAA GCC TTA GAG TCT CCT GAG CAT TGC     144
Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu Glu Ser Pro Glu His Cys
        35                  40                  45

TCA CCT CAC CAC ACC GCA CTC AGG CAA GCC ATT CTC TGC TGG GGG GAA     192
Ser Pro His His Thr Ala Leu Arg Gln Ala Ile Leu Cys Trp Gly Glu
    50                  55                  60

TTG ATG ACT CTA GCT ACC TGG GTG GGT AAT AAT TTG GAG GAT CCA GCA     240
Leu Met Thr Leu Ala Thr Trp Val Gly Asn Asn Leu Glu Asp Pro Ala
65                  70                  75                  80

TCA AGG GAT CTA GTA GTC AAT TAT GTT AAT ACT AAC ATG GGT TTA AAA     288
Ser Arg Asp Leu Val Val Asn Tyr Val Asn Thr Asn Met Gly Leu Lys
                85                  90                  95

ATT AGG CAA CTA TTG TGG TTT CAT ATA TCT TGC CTT ACT TTT GGA AGA     336
Ile Arg Gln Leu Leu Trp Phe His Ile Ser Cys Leu Thr Phe Gly Arg
            100                 105                 110

GAG ACT GTA CTT GAA TAT TTG GTA TCT TTC GGA GTG TGG ATT CGC ACT     384
Glu Thr Val Leu Glu Tyr Leu Val Ser Phe Gly Val Trp Ile Arg Thr
        115                 120                 125

CCT CCA GCC TAT AGA CCA CCA AAT GCC CCT ATC TTA TCA ACA CTT CCG     432
Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro
    130                 135                 140

GAA ACT ACT GTT GTT AGA CGA CGG GAC CGA GGC AGG TCC CCT AGA AGA     480
Glu Thr Thr Val Val Arg Arg Arg Asp Arg Gly Arg Ser Pro Arg Arg
145                 150                 155                 160

AGA ACT CCC TCG CCT CGC AGA CGC AGA TCT CAA TCT CGG GAA TCT CAA     528
Arg Thr Pro Ser Pro Arg Arg Arg Arg Ser Gln Ser Arg Glu Ser Gln
                165                 170                 175

TGT TAG                                                            534
Cys
```

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 177 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```
Met Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ala Thr Val Glu Leu Leu
 1               5                  10                  15

Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser Val Arg Asp Leu Leu Asp
            20                  25                  30

Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu Glu Ser Pro Glu His Cys
        35                  40                  45

Ser Pro His His Thr Ala Leu Arg Gln Ala Ile Leu Cys Trp Gly Glu
    50                  55                  60

Leu Met Thr Leu Ala Thr Trp Val Gly Asn Asn Leu Glu Asp Pro Ala
65                  70                  75                  80

Ser Arg Asp Leu Val Val Asn Tyr Val Asn Thr Asn Met Gly Leu Lys
                85                  90                  95

Ile Arg Gln Leu Leu Trp Phe His Ile Ser Cys Leu Thr Phe Gly Arg
            100                 105                 110

Glu Thr Val Leu Glu Tyr Leu Val Ser Phe Gly Val Trp Ile Arg Thr
        115                 120                 125

Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro
    130                 135                 140

Glu Thr Thr Val Val Arg Arg Arg Asp Arg Gly Arg Ser Pro Arg Arg
145                 150                 155                 160

Arg Thr Pro Ser Pro Arg Arg Arg Arg Ser Gln Ser Arg Glu Ser Gln
                165                 170                 175

Cys
```

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 558 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..558

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```
ATG GAC ATT GAC CCT TAT AAA GAA TTT GGA GCT ACT GTG GAG TTA CTC        48
Met Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ala Thr Val Glu Leu Leu
 1               5                  10                  15

TCG TTT TTG CCT TCT GAC TTC TTT CCT TCC GTC AGA GAT CTC CTA GAC        96
Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser Val Arg Asp Leu Leu Asp
            20                  25                  30

ACC GCC TCA GCT CTG TAT CGG GAA GCC TTA GAG TCT CCT GAG CAT TGC       144
Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu Glu Ser Pro Glu His Cys
        35                  40                  45
```

```
TCA CCT CAC CAC ACC GCA CTC AGG CAA GCC ATT CTC TGC TGG GGG GAA        192
Ser Pro His His Thr Ala Leu Arg Gln Ala Ile Leu Cys Trp Gly Glu
     50                  55                  60

TTG ATG ACT CTA GCT ACC TGG GTG GGT AAT AAT TTG GAG GAT CCA GCA        240
Leu Met Thr Leu Ala Thr Trp Val Gly Asn Asn Leu Glu Asp Pro Ala
 65                  70                  75                  80

TCA AGG GAT CTA GTA GTC AAT TAT GTT AAT ACT AAC ATG GGT TTA AAA        288
Ser Arg Asp Leu Val Val Asn Tyr Val Asn Thr Asn Met Gly Leu Lys
                85                  90                  95

ATT AGG CAA CTA TTG TGG TTT CAT ATA TCT TGC CTT ACT TTT GGA AGA        336
Ile Arg Gln Leu Leu Trp Phe His Ile Ser Cys Leu Thr Phe Gly Arg
            100                 105                 110

GAG ACT GTA CTT GAA TAT TTG GTA TCT TTC GGA GTG TGG ATT CGC ACT        384
Glu Thr Val Leu Glu Tyr Leu Val Ser Phe Gly Val Trp Ile Arg Thr
        115                 120                 125

CCT CCA GCC TAT AGA CCA CCA AAT GCC CCT ATC TTA TCA ACA CTT CCG        432
Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro
    130                 135                 140

GAA ACT ACT GTT GTT AGA CGA CGG GAC CGA GGC AGG TCC CCT AGA AGA        480
Glu Thr Thr Val Val Arg Arg Arg Asp Arg Gly Arg Ser Pro Arg Arg
145                 150                 155                 160

AGA ACT CCC TCG CCT CGC AGA CGC AGA TCC CAA TCG CCG CGT CGC AGA        528
Arg Thr Pro Ser Pro Arg Arg Arg Arg Ser Gln Ser Pro Arg Arg Arg
                165                 170                 175

CGA TCT CAA TCT CGG GAA TCT CAA TGT TAG                              558
Arg Ser Gln Ser Arg Glu Ser Gln Cys
            180                 185
```

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 185 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```
Met Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ala Thr Val Glu Leu Leu
 1               5                  10                  15

Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser Val Arg Asp Leu Leu Asp
            20                  25                  30

Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu Glu Ser Pro Glu His Cys
        35                  40                  45

Ser Pro His His Thr Ala Leu Arg Gln Ala Ile Leu Cys Trp Gly Glu
    50                  55                  60

Leu Met Thr Leu Ala Thr Trp Val Gly Asn Asn Leu Glu Asp Pro Ala
65                  70                  75                  80

Ser Arg Asp Leu Val Val Asn Tyr Val Asn Thr Asn Met Gly Leu Lys
                85                  90                  95

Ile Arg Gln Leu Leu Trp Phe His Ile Ser Cys Leu Thr Phe Gly Arg
            100                 105                 110

Glu Thr Val Leu Glu Tyr Leu Val Ser Phe Gly Val Trp Ile Arg Thr
```

115 120 125

Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro
130 135 140

Glu Thr Thr Val Val Arg Arg Arg Asp Arg Gly Arg Ser Pro Arg Arg
145 150 155 160

Arg Thr Pro Ser Pro Arg Arg Arg Arg Ser Gln Ser Pro Arg Arg Arg
165 170 175

Arg Ser Gln Ser Arg Glu Ser Gln Cys
180 185

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GATCCCAATC GCGGCGTCGC AGAC 24

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GATCGTCTGC GACGCGGCGA TTGG 24

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CCCTCGAGGG 10

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CTCACTCGAG GGAATAAGCT 20

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TTGTTAGGCC TTAAATGTAT 20

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GAGCTCGGTA CCACCATGAA GTGGA 25

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GAGCAGGGGT CCTAGGAATC 20

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

TCACTATCCG GAATCAGCTT 20

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GGTGAATTCT GGCCCGAATG 20

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 627 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..627

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

```
ATG GAC ATT GAC CCT TAT AAA GAA TTT GGA GCT ACT GTG GAG TTA CTC      48
Met Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ala Thr Val Glu Leu Leu
 1               5                  10                  15

TCG TTT TTG CCT TCT GAC TTC TTT CCT TCC GTC AGA GAT CTC CTA GAC      96
Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser Val Arg Asp Leu Leu Asp
            20                  25                  30

ACC GCC TCA GCT CTG TAT CGG GAA GCC TTA GAG TCT CCT GAG CAT TGC     144
Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu Glu Ser Pro Glu His Cys
        35                  40                  45

TCA CCT CAC CAC ACC GCA CTC AGG CAA GCC ATT CTC TGC TGG GGG GAA     192
Ser Pro His His Thr Ala Leu Arg Gln Ala Ile Leu Cys Trp Gly Glu
    50                  55                  60

TTG ATG ACT CTA GCT ACC TGG GTG GGT AAT AAT TTG GAG GAT CCA GCA     240
Leu Met Thr Leu Ala Thr Trp Val Gly Asn Asn Leu Glu Asp Pro Ala
65                  70                  75                  80

TCA AGG GAT CTA GTA GTC AAT TAT GTT AAT ACT AAC ATG GGT TTA AAA     288
Ser Arg Asp Leu Val Val Asn Tyr Val Asn Thr Asn Met Gly Leu Lys
                85                  90                  95

ATT AGG CAA CTA TTG TGG TTT CAT ATA TCT TGC CTT ACT TTT GGA AGA     336
Ile Arg Gln Leu Leu Trp Phe His Ile Ser Cys Leu Thr Phe Gly Arg
            100                 105                 110

GAG ACT GTA CTT GAA TAT TTG GTA TCT TTC GGA GTG TGG ATT CGC ACT     384
Glu Thr Val Leu Glu Tyr Leu Val Ser Phe Gly Val Trp Ile Arg Thr
        115                 120                 125

CCT CCA GCC TAT AGA CCA CCA AAT GCC CCT ATC TTA TCA ACA CTT CCG     432
Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro
```

```
     130                 135                 140

GAA TCA GCT TGC ATG GGG ACG AAT CTT TCT GTT CCC AAT CCT CTG GGA        480
Glu Ser Ala Cys Met Gly Thr Asn Leu Ser Val Pro Asn Pro Leu Gly
145                 150                 155                 160

TTC TTT CCC GAT CAT CAG TTG GAC CCT GCA TTC GGA GCC AAC TCA AAC        528
Phe Phe Pro Asp His Gln Leu Asp Pro Ala Phe Gly Ala Asn Ser Asn
                165                 170                 175

AAT CCA GAT TGG GAC TTC AAC CCC ATC AAG GAC CAC TGG CCA GCA GCC        576
Asn Pro Asp Trp Asp Phe Asn Pro Ile Lys Asp His Trp Pro Ala Ala
            180                 185                 190

AAC CAG GTA GGA GTG GGA GCA TTC GGG CCA GAA TTC AGG CCT ACT AGT        624
Asn Gln Val Gly Val Gly Ala Phe Gly Pro Glu Phe Arg Pro Thr Ser
        195                 200                 205

TAA                                                                    627
```

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 208 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

```
Met Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ala Thr Val Glu Leu Leu
  1               5                  10                  15

Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser Val Arg Asp Leu Leu Asp
             20                  25                  30

Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu Glu Ser Pro Glu His Cys
         35                  40                  45

Ser Pro His His Thr Ala Leu Arg Gln Ala Ile Leu Cys Trp Gly Glu
     50                  55                  60

Leu Met Thr Leu Ala Thr Trp Val Gly Asn Asn Leu Glu Asp Pro Ala
 65                  70                  75                  80

Ser Arg Asp Leu Val Val Asn Tyr Val Asn Thr Asn Met Gly Leu Lys
                 85                  90                  95

Ile Arg Gln Leu Leu Trp Phe His Ile Ser Cys Leu Thr Phe Gly Arg
            100                 105                 110

Glu Thr Val Leu Glu Tyr Leu Val Ser Phe Gly Val Trp Ile Arg Thr
        115                 120                 125

Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro
    130                 135                 140

Glu Ser Ala Cys Met Gly Thr Asn Leu Ser Val Pro Asn Pro Leu Gly
145                 150                 155                 160

Phe Phe Pro Asp His Gln Leu Asp Pro Ala Phe Gly Ala Asn Ser Asn
                165                 170                 175

Asn Pro Asp Trp Asp Phe Asn Pro Ile Lys Asp His Trp Pro Ala Ala
            180                 185                 190
```

Asn Gln Val Gly Val Gly Ala Phe Gly Pro Glu Phe Arg Pro Thr Ser
195 200 205

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

CATCCTCCGG ACATGCAGTG 20

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GTCCTAGGAA TTCTGATGTG 20

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 639 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..639

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

ATG GAC ATT GAC CCT TAT AAA GAA TTT GGA GCT ACT GTG GAG TTA CTC 48
Met Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ala Thr Val Glu Leu Leu
1 5 10 15

TCG TTT TTG CCT TCT GAC TTC TTT CCT TCC GTC AGA GAT CTC CTA GAC 96
Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser Val Arg Asp Leu Leu Asp
20 25 30

ACC GCC TCA GCT CTG TAT CGG GAA GCC TTA GAG TCT CCT GAG CAT TGC 144
Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu Glu Ser Pro Glu His Cys
35 40 45

TCA CCT CAC CAC ACC GCA CTC AGG CAA GCC ATT CTC TGC TGG GGG GAA 192

```
Ser Pro His His Thr Ala Leu Arg Gln Ala Ile Leu Cys Trp Gly Glu
 50                  55                  60

TTG ATG ACT CTA GCT ACC TGG GTG GGT AAT AAT TTG GAG GAT CCA GCA      240
Leu Met Thr Leu Ala Thr Trp Val Gly Asn Asn Leu Glu Asp Pro Ala
 65                  70                  75                  80

TCA AGG GAT CTA GTA GTC AAT TAT GTT AAT ACT AAC ATG GGT TTA AAA      288
Ser Arg Asp Leu Val Val Asn Tyr Val Asn Thr Asn Met Gly Leu Lys
                 85                  90                  95

ATT AGG CAA CTA TTG TGG TTT CAT ATA TCT TGC CTT ACT TTT GGA AGA      336
Ile Arg Gln Leu Leu Trp Phe His Ile Ser Cys Leu Thr Phe Gly Arg
            100                 105                 110

GAG ACT GTA CTT GAA TAT TTG GTA TCT TTC GGA GTG TGG ATT CGC ACT      384
Glu Thr Val Leu Glu Tyr Leu Val Ser Phe Gly Val Trp Ile Arg Thr
        115                 120                 125

CCT CCA GCC TAT AGA CCA CCA AAT GCC CCT ATC TTA TCA ACA CTT CCG      432
Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro
    130                 135                 140

GAC ATG CAG TGG AAC TCC ACT GCC TTC CAC CAA GCT CTG CAG GAT CCC      480
Asp Met Gln Trp Asn Ser Thr Ala Phe His Gln Ala Leu Gln Asp Pro
145                 150                 155                 160

AGA GTC AGG GGT CTG TAT TTT CCT GCT GGT GGC TCC AGT TCA GGA ACA      528
Arg Val Arg Gly Leu Tyr Phe Pro Ala Gly Gly Ser Ser Ser Gly Thr
                165                 170                 175

GTA AAC CCT GCT CCG AAT ATT GCC TCT CAC ATC TCG TCA ATC TCC GCG      576
Val Asn Pro Ala Pro Asn Ile Ala Ser His Ile Ser Ser Ile Ser Ala
            180                 185                 190

AGG ACT GGG GAC CCT GTG ACG AAC ATG GAG AAC ATC ACA TCA GAA TTC      624
Arg Thr Gly Asp Pro Val Thr Asn Met Glu Asn Ile Thr Ser Glu Phe
        195                 200                 205

AGG CCT ACT AGT TAA                                                  639
Arg Pro Thr Ser
    210
```

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 212 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

```
Met Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ala Thr Val Glu Leu Leu
  1                 5                  10                  15

Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser Val Arg Asp Leu Leu Asp
             20                  25                  30

Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu Glu Ser Pro Glu His Cys
         35                  40                  45

Ser Pro His His Thr Ala Leu Arg Gln Ala Ile Leu Cys Trp Gly Glu
     50                  55                  60

Leu Met Thr Leu Ala Thr Trp Val Gly Asn Asn Leu Glu Asp Pro Ala
 65                  70                  75                  80
```

Ser Arg Asp Leu Val Val Asn Tyr Val Asn Thr Asn Met Gly Leu Lys
85 90 95

Ile Arg Gln Leu Leu Trp Phe His Ile Ser Cys Leu Thr Phe Gly Arg
100 105 110

Glu Thr Val Leu Glu Tyr Leu Val Ser Phe Gly Val Trp Ile Arg Thr
115 120 125

Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro
130 135 140

Asp Met Gln Trp Asn Ser Thr Ala Phe His Gln Ala Leu Gln Asp Pro
145 150 155 160

Arg Val Arg Gly Leu Tyr Phe Pro Ala Gly Gly Ser Ser Ser Gly Thr
165 170 175

Val Asn Pro Ala Pro Asn Ile Ala Ser His Ile Ser Ser Ile Ser Ala
180 185 190

Arg Thr Gly Asp Pro Val Thr Asn Met Glu Asn Ile Thr Ser Glu Phe
195 200 205

Arg Pro Thr Ser
210

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GTATGTTTCC GGATTGTCCT 20

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

TGAGGCCGAA TTCCATAGGT 20

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 630 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..624

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

```
ATG GAC ATT GAC CCT TAT AAA GAA TTT GGA GCT ACT GTG GAG TTA CTC          48
Met Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ala Thr Val Glu Leu Leu
 1               5                  10                  15

TCG TTT TTG CCT TCT GAC TTC TTT CCT TCC GTC AGA GAT CTC CTA GAC          96
Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser Val Arg Asp Leu Leu Asp
            20                  25                  30

ACC GCC TCA GCT CTG TAT CGG GAA GCC TTA GAG TCT CCT GAG CAT TGC         144
Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu Glu Ser Pro Glu His Cys
        35                  40                  45

TCA CCT CAC CAC ACC GCA CTC AGG CAA GCC ATT CTC TGC TGG GGG GAA         192
Ser Pro His His Thr Ala Leu Arg Gln Ala Ile Leu Cys Trp Gly Glu
    50                  55                  60

TTG ATG ACT CTA GCT ACC TGG GTG GGT AAT AAT TTG GAG GAT CCA GCA         240
Leu Met Thr Leu Ala Thr Trp Val Gly Asn Asn Leu Glu Asp Pro Ala
65                  70                  75                  80

TCA AGG GAT CTA GTA GTC AAT TAT GTT AAT ACT AAC ATG GGT TTA AAA         288
Ser Arg Asp Leu Val Val Asn Tyr Val Asn Thr Asn Met Gly Leu Lys
                85                  90                  95

ATT AGG CAA CTA TTG TGG TTT CAT ATA TCT TGC CTT ACT TTT GGA AGA         336
Ile Arg Gln Leu Leu Trp Phe His Ile Ser Cys Leu Thr Phe Gly Arg
            100                 105                 110

GAG ACT GTA CTT GAA TAT TTG GTA TCT TTC GGA GTG TGG ATT CGC ACT         384
Glu Thr Val Leu Glu Tyr Leu Val Ser Phe Gly Val Trp Ile Arg Thr
        115                 120                 125

CCT CCA GCC TAT AGA CCA CCA AAT GCC CCT ATC TTA TCA ACA CTT CCG         432
Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro
    130                 135                 140

GAT TGT CCT CTA ATT CCA GGA TCA ACA ACA ACC AGT ACG GGA CCA TGC         480
Asp Cys Pro Leu Ile Pro Gly Ser Thr Thr Thr Ser Thr Gly Pro Cys
145                 150                 155                 160

AAA ACC TGC ACG ACT CCT GCT CAA GGC AAC TCT ATG TTT CCC TCA TGT         528
Lys Thr Cys Thr Thr Pro Ala Gln Gly Asn Ser Met Phe Pro Ser Cys
                165                 170                 175

TGC TGT ACA AAA CCT ACG GAT GGA AAT TGC ACC TGT ATT CCC ATC CCA         576
Cys Cys Thr Lys Pro Thr Asp Gly Asn Cys Thr Cys Ile Pro Ile Pro
            180                 185                 190

TCG TCT TGG GCT TTC GCA AAA TAC CTA TGG AAT TCA GGC CTA CTA GTT         624
Ser Ser Trp Ala Phe Ala Lys Tyr Leu Trp Asn Ser Gly Leu Leu Val
        195                 200                 205

AAG TAA
Lys                                                             630
```

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 209 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Met Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ala Thr Val Glu Leu Leu
1 5 10 15

Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser Val Arg Asp Leu Leu Asp
20 25 30

Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu Glu Ser Pro Glu His Cys
35 40 45

Ser Pro His His Thr Ala Leu Arg Gln Ala Ile Leu Cys Trp Gly Glu
50 55 60

Leu Met Thr Leu Ala Thr Trp Val Gly Asn Asn Leu Glu Asp Pro Ala
65 70 75 80

Ser Arg Asp Leu Val Val Asn Tyr Val Asn Thr Asn Met Gly Leu Lys
85 90 95

Ile Arg Gln Leu Leu Trp Phe His Ile Ser Cys Leu Thr Phe Gly Arg
100 105 110

Glu Thr Val Leu Glu Tyr Leu Val Ser Phe Gly Val Trp Ile Arg Thr
115 120 125

Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro
130 135 140

Asp Cys Pro Leu Ile Pro Gly Ser Thr Thr Thr Ser Thr Gly Pro Cys
145 150 155 160

Lys Thr Cys Thr Thr Pro Ala Gln Gly Asn Ser Met Phe Pro Ser Cys
165 170 175

Cys Cys Thr Lys Pro Thr Asp Gly Asn Cys Thr Cys Ile Pro Ile Pro
180 185 190

Ser Ser Trp Ala Phe Ala Lys Tyr Leu Trp Asn Ser Gly Leu Leu Val
195 200 205

Lys

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

CGAGGAATTT ATTTATAGC 19

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

CTCACTCGAG GGAATAAGCT 20

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

TTGTTAGGCC TTAAATGTAT 20

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

GTGGGTAAGC TTCTCGATGT 20

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

AGTTTCCAAG CTTATGAG 18

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

GCGCCATGGA CATTGACCCT TATA 24

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

CCCTGATCAC TAACATTGAG ATTCCCGA 28

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

ATTCTCGAGG GCATGGGGAC G 21

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

ACTCCATGGC CTGAGGATGA 20

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 858 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..858

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

ATG GGG ACG AAT CTT TCT GTT CCC AAT CCT CTG GGA TTC TTT CCC GAT 48
Met Gly Thr Asn Leu Ser Val Pro Asn Pro Leu Gly Phe Phe Pro Asp

```
1               5                  10                  15

CAT CAG TTG GAC CCT GCA TTC GGA GCC AAC TCA AAC AAT CCA GAT TGG        96
His Gln Leu Asp Pro Ala Phe Gly Ala Asn Ser Asn Asn Pro Asp Trp
            20                  25                  30

GAC TTC AAC CCC ATC AAG GAC CAC TGG CCA GCA GCC AAC CAG GTA GGA       144
Asp Phe Asn Pro Ile Lys Asp His Trp Pro Ala Ala Asn Gln Val Gly
        35                  40                  45

GTG GGA GCA TTC GGG CCA GGG TTC ACC CCT CCA CAC GGC GGT GTT TTG       192
Val Gly Ala Phe Gly Pro Gly Phe Thr Pro Pro His Gly Gly Val Leu
    50                  55                  60

GGG TGG AGC CCT CAG GCT CAG GGC ATA TTG ACC ACA GTG TCA ACA ATT       240
Gly Trp Ser Pro Gln Ala Gln Gly Ile Leu Thr Thr Val Ser Thr Ile
65                  70                  75                  80

CCT CCT CCT GCC TCC ACC AAT CGG CAG TCA GGA AGG CAG CCT ACT CCC       288
Pro Pro Pro Ala Ser Thr Asn Arg Gln Ser Gly Arg Gln Pro Thr Pro
                85                  90                  95

ATC TCT CCA CCT CTA AGA GAC AGT CAT CCT CAG GCC ATG GAC ATT GAC       336
Ile Ser Pro Pro Leu Arg Asp Ser His Pro Gln Ala Met Asp Ile Asp
            100                 105                 110

CCT TAT AAA GAA TTT GGA GCT ACT GTG GAG TTA CTC TCG TTT TTG CCT       384
Pro Tyr Lys Glu Phe Gly Ala Thr Val Glu Leu Leu Ser Phe Leu Pro
        115                 120                 125

TCT GAC TTC TTT CCT TCC GTC AGA GAT CTC CTA GAC ACC GCC TCA GCT       432
Ser Asp Phe Phe Pro Ser Val Arg Asp Leu Leu Asp Thr Ala Ser Ala
    130                 135                 140

CTG TAT CGG GAA GCC TTA GAG TCT CCT GAG CAT TGC TCA CCT CAC CAC       480
Leu Tyr Arg Glu Ala Leu Glu Ser Pro Glu His Cys Ser Pro His His
145                 150                 155                 160

ACC GCA CTC AGG CAA GCC ATT CTC TGC TGG GGG GAA TTG ATG ACT CTA       528
Thr Ala Leu Arg Gln Ala Ile Leu Cys Trp Gly Glu Leu Met Thr Leu
                165                 170                 175

GCT ACC TGG GTG GGT AAT AAT TTG GAG GAT CCA GCA TCA AGG GAT CTA       576
Ala Thr Trp Val Gly Asn Asn Leu Glu Asp Pro Ala Ser Arg Asp Leu
            180                 185                 190

GTA GTC AAT TAT GTT AAT ACT AAC ATG GGT TTA AAA ATT AGG CAA CTA       624
Val Val Asn Tyr Val Asn Thr Asn Met Gly Leu Lys Ile Arg Gln Leu
        195                 200                 205

TTG TGG TTT CAT ATA TCT TGC CTT ACT TTT GGA AGA GAG ACT GTA CTT       672
Leu Trp Phe His Ile Ser Cys Leu Thr Phe Gly Arg Glu Thr Val Leu
    210                 215                 220

GAA TAT TTG GTA TCT TTC GGA GTG TGG ATT CGC ACT CCT CCA GCC TAT       720
Glu Tyr Leu Val Ser Phe Gly Val Trp Ile Arg Thr Pro Pro Ala Tyr
225                 230                 235                 240

AGA CCA CCA AAT GCC CCT ATC TTA TCA ACA CTT CCG GAA ACT ACT GTT       768
Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro Glu Thr Thr Val
                245                 250                 255

GTT AGA CGA CGG GAC CGA GGC AGG TCC CCT AGA AGA AGA ACT CCC TCG       816
Val Arg Arg Arg Asp Arg Gly Arg Ser Pro Arg Arg Arg Thr Pro Ser
            260                 265                 270

CCT CGC AGA CGC AGA TCT CAA TCT CGG GAA TCT CAA TGT TAG               858
Pro Arg Arg Arg Arg Ser Gln Ser Arg Glu Ser Gln Cys
        275                 280                 285
```

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 285 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

```
Met Gly Thr Asn Leu Ser Val Pro Asn Pro Leu Gly Phe Phe Pro Asp
 1               5                  10                  15

His Gln Leu Asp Pro Ala Phe Gly Ala Asn Ser Asn Asn Pro Asp Trp
            20                  25                  30

Asp Phe Asn Pro Ile Lys Asp His Trp Pro Ala Ala Asn Gln Val Gly
        35                  40                  45

Val Gly Ala Phe Gly Pro Gly Phe Thr Pro Pro His Gly Gly Val Leu
    50                  55                  60

Gly Trp Ser Pro Gln Ala Gln Gly Ile Leu Thr Thr Val Ser Thr Ile
65                  70                  75                  80

Pro Pro Pro Ala Ser Thr Asn Arg Gln Ser Gly Arg Gln Pro Thr Pro
                85                  90                  95

Ile Ser Pro Pro Leu Arg Asp Ser His Pro Gln Ala Met Asp Ile Asp
            100                 105                 110

Pro Tyr Lys Glu Phe Gly Ala Thr Val Glu Leu Leu Ser Phe Leu Pro
        115                 120                 125

Ser Asp Phe Phe Pro Ser Val Arg Asp Leu Leu Asp Thr Ala Ser Ala
    130                 135                 140

Leu Tyr Arg Glu Ala Leu Glu Ser Pro Glu His Cys Ser Pro His His
145                 150                 155                 160

Thr Ala Leu Arg Gln Ala Ile Leu Cys Trp Gly Glu Leu Met Thr Leu
                165                 170                 175

Ala Thr Trp Val Gly Asn Asn Leu Glu Asp Pro Ala Ser Arg Asp Leu
            180                 185                 190

Val Val Asn Tyr Val Asn Thr Asn Met Gly Leu Lys Ile Arg Gln Leu
        195                 200                 205

Leu Trp Phe His Ile Ser Cys Leu Thr Phe Gly Arg Glu Thr Val Leu
    210                 215                 220

Glu Tyr Leu Val Ser Phe Gly Val Trp Ile Arg Thr Pro Pro Ala Tyr
225                 230                 235                 240

Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro Glu Thr Thr Val
                245                 250                 255

Val Arg Arg Arg Asp Arg Gly Arg Ser Pro Arg Arg Arg Thr Pro Ser
            260                 265                 270

Pro Arg Arg Arg Arg Ser Gln Ser Arg Glu Ser Gln Cys
        275                 280                 285
```

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 nucleotides

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

GCGCCATGGA CATTGACCCT TATA 24

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

AGTGAAGCTT CCCACCTT 18

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 882 nucleotides

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..882

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

```
ATG GGG ACG AAT CTT TCT GTT CCC AAT CCT CTG GGA TTC TTT CCC GAT      48
Met Gly Thr Asn Leu Ser Val Pro Asn Pro Leu Gly Phe Phe Pro Asp
 1               5                  10                  15

CAT CAG TTG GAC CCT GCA TTC GGA GCC AAC TCA AAC AAT CCA GAT TGG      96
His Gln Leu Asp Pro Ala Phe Gly Ala Asn Ser Asn Asn Pro Asp Trp
            20                  25                  30

GAC TTC AAC CCC ATC AAG GAC CAC TGG CCA GCA GCC AAC CAG GTA GGA     144
Asp Phe Asn Pro Ile Lys Asp His Trp Pro Ala Ala Asn Gln Val Gly
        35                  40                  45

GTG GGA GCA TTC GGG CCA GGG TTC ACC CCT CCA CAC GGC GGT GTT TTG     192
Val Gly Ala Phe Gly Pro Gly Phe Thr Pro Pro His Gly Gly Val Leu
    50                  55                  60

GGG TGG AGC CCT CAG GCT CAG GGC ATA TTG ACC ACA GTG TCA ACA ATT     240
Gly Trp Ser Pro Gln Ala Gln Gly Ile Leu Thr Thr Val Ser Thr Ile
65                  70                  75                  80

CCT CCT CCT GCC TCC ACC AAT CGG CAG TCA GGA AGG CAG CCT ACT CCC     288
Pro Pro Pro Ala Ser Thr Asn Arg Gln Ser Gly Arg Gln Pro Thr Pro
                85                  90                  95
```

```
ATC TCT CCA CCT CTA AGA GAC AGT CAT CCT CAG GCC ATG GAC ATT GAC          336
Ile Ser Pro Pro Leu Arg Asp Ser His Pro Gln Ala Met Asp Ile Asp
            100                 105                 110

CCT TAT AAA GAA TTT GGA GCT ACT GTG GAG TTA CTC TCG TTT TTG CCT          384
Pro Tyr Lys Glu Phe Gly Ala Thr Val Glu Leu Leu Ser Phe Leu Pro
        115                 120                 125

TCT GAC TTC TTT CCT TCC GTC AGA GAT CTC CTA GAC ACC GCC TCA GCT          432
Ser Asp Phe Phe Pro Ser Val Arg Asp Leu Leu Asp Thr Ala Ser Ala
    130                 135                 140

CTG TAT CGG GAA GCC TTA GAG TCT CCT GAG CAT TGC TCA CCT CAC CAC          480
Leu Tyr Arg Glu Ala Leu Glu Ser Pro Glu His Cys Ser Pro His His
145                 150                 155                 160

ACC GCA CTC AGG CAA GCC ATT CTC TGC TGG GGG GAA TTG ATG ACT CTA          528
Thr Ala Leu Arg Gln Ala Ile Leu Cys Trp Gly Glu Leu Met Thr Leu
                165                 170                 175

GCT ACC TGG GTG GGT AAT AAT TTG GAG GAT CCA GCA TCA AGG GAT CTA          576
Ala Thr Trp Val Gly Asn Asn Leu Glu Asp Pro Ala Ser Arg Asp Leu
            180                 185                 190

GTA GTC AAT TAT GTT AAT ACT AAC ATG GGT TTA AAA ATT AGG CAA CTA          624
Val Val Asn Tyr Val Asn Thr Asn Met Gly Leu Lys Ile Arg Gln Leu
        195                 200                 205

TTG TGG TTT CAT ATA TCT TGC CTT ACT TTT GGA AGA GAG ACT GTA CTT          672
Leu Trp Phe His Ile Ser Cys Leu Thr Phe Gly Arg Glu Thr Val Leu
    210                 215                 220

GAA TAT TTG GTA TCT TTC GGA GTG TGG ATT CGC ACT CCT CCA GCC TAT          720
Glu Tyr Leu Val Ser Phe Gly Val Trp Ile Arg Thr Pro Pro Ala Tyr
225                 230                 235                 240

AGA CCA CCA AAT GCC CCT ATC TTA TCA ACA CTT CCG GAA ACT ACT GTT          768
Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro Glu Thr Thr Val
                245                 250                 255

GTT AGA CGA CGG GAC CGA GGC AGG TCC CCT AGA AGA AGA ACT CCC TCG          816
Val Arg Arg Arg Asp Arg Gly Arg Ser Pro Arg Arg Arg Thr Pro Ser
            260                 265                 270

CCT CGC AGA CGC AGA TCC CAA TCG CCG CGT CGC AGA CGA TCT CAA TCT          864
Pro Arg Arg Arg Arg Ser Gln Ser Pro Arg Arg Arg Arg Ser Gln Ser
        275                 280                 285

CGG GAA TCT CAA TGT TAG                                                 882
Arg Glu Ser Gln Cys
    290
```

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 293 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

```
Met Gly Thr Asn Leu Ser Val Pro Asn Pro Leu Gly Phe Phe Pro Asp
 1               5                  10                  15

His Gln Leu Asp Pro Ala Phe Gly Ala Asn Ser Asn Asn Pro Asp Trp
            20                  25                  30
```

```
Asp Phe Asn Pro Ile Lys Asp His Trp Pro Ala Ala Asn Gln Val Gly
 35                  40                  45

Val Gly Ala Phe Gly Pro Gly Phe Thr Pro Pro His Gly Gly Val Leu
 50                  55                  60

Gly Trp Ser Pro Gln Ala Gln Gly Ile Leu Thr Thr Val Ser Thr Ile
 65                  70                  75                  80

Pro Pro Pro Ala Ser Thr Asn Arg Gln Ser Gly Arg Gln Pro Thr Pro
                 85                  90                  95

Ile Ser Pro Pro Leu Arg Asp Ser His Pro Gln Ala Met Asp Ile Asp
            100                 105                 110

Pro Tyr Lys Glu Phe Gly Ala Thr Val Glu Leu Leu Ser Phe Leu Pro
        115                 120                 125

Ser Asp Phe Phe Pro Ser Val Arg Asp Leu Leu Asp Thr Ala Ser Ala
    130                 135                 140

Leu Tyr Arg Glu Ala Leu Glu Ser Pro Glu His Cys Ser Pro His His
145                 150                 155                 160

Thr Ala Leu Arg Gln Ala Ile Leu Cys Trp Gly Glu Leu Met Thr Leu
                165                 170                 175

Ala Thr Trp Val Gly Asn Asn Leu Glu Asp Pro Ala Ser Arg Asp Leu
            180                 185                 190

Val Val Asn Tyr Val Asn Thr Asn Met Gly Leu Lys Ile Arg Gln Leu
        195                 200                 205

Leu Trp Phe His Ile Ser Cys Leu Thr Phe Gly Arg Glu Thr Val Leu
    210                 215                 220

Glu Tyr Leu Val Ser Phe Gly Val Trp Ile Arg Thr Pro Pro Ala Tyr
225                 230                 235                 240

Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro Glu Thr Thr Val
                245                 250                 255

Val Arg Arg Arg Asp Arg Gly Arg Ser Pro Arg Arg Arg Thr Pro Ser
            260                 265                 270

Pro Arg Arg Arg Arg Ser Gln Ser Pro Arg Arg Arg Arg Ser Gln Ser
        275                 280                 285

Arg Glu Ser Gln Cys
    290
```

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

GTTGGTAGAA TTCCAATTAT 20

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

AAGAATGATC GATACAGTTT 20

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

GTACCCCGAA TTCATACTTA 20

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

AATAATATAT CGATAATTGT 20

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

GTTGGAATTC CGCTACTGAT 20

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

TAACCAAGTA TCGATATAAT 20

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

Arg Pro Thr Ser
1

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

Asn Ser Gly Leu Leu Val Lys
1 5

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

CTCGATGTCG ACTAGCCATA 20

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:
AAGTTGTCGA CCTTATGAGT

International Application No: PCT/ /

MICROORGANISMS

Optional Sheet in connection with the microorganism referred to on page 107, lines 20-35 of the description

A. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet

Name of depositary institution

American Type Culture Collection

Address of depositary institution (including postal code and country)

12301 Parklawn Drive
Rockville, MD 20852
US

Date of deposit November 12, 1992 Accession Number 69124

B. ADDITIONAL INDICATIONS (leave blank if not applicable). This information is continued on a separate attached sheet

C. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not all designated States)

D. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)

The indications listed below will be submitted to the International Bureau later (Specify the general nature of the indications e.g., "Accession Number of Deposit")

E. ☐ This sheet was received with the International application when filed (to be checked by the receiving Office)

(Authorized Officer)

☐ The date of receipt (from the applicant) by the International Bureau

was

(Authorized Officer)

Form PCT/RO/134 (January 1981)

WHAT IS CLAIMED IS:

1. A recombinant replicable vaccinia virus, the genome of which comprises:

(a) a first nucleotide sequence encoding at least one epitope of a hepatitis B virus surface antigen;

(b) a first promoter sequence operably linked to said first nucleotide sequence;

(c) a second nucleotide sequence encoding at least one epitope of a hepatitis B virus core antigen; and

(d) a second promoter sequence operably linked to said second nucleotide sequence;

whereby the first and second nucleotide sequences are expressed by the virus in a suitable host.

2. The virus of claim 1 in which the first nucleotide sequence encodes an amino acid sequence comprising a protein selected from the group consisting of a hepatitis B virus S, MS, LS, preS2, preS1, derivatives of any of the foregoing displaying the antigenicity of any of the foregoing, and fusion proteins consisting of combinations of S, MS, LS, preS2, preS1, and derivatives of any of the foregoing displaying the antigenicity of any of the foregoing; and the second nucleotide sequence encodes an amino acid sequence comprising a protein selected from the group consisting of core antigen, e antigen, and derivatives thereof displaying the antigenicity of core antigen or e antigen.

3. The virus of claim 1 which is infectious in the host without causing significant disease.

4. A recombinant replicable vaccinia virus, the genome of which comprises:

(a) a first nucleotide sequence encoding hepatitis B virus LS, MS, S, or a derivative of any of the foregoing displaying the antigenicity of any of the foregoing;

(b) a first promoter sequence operably linked to said first nucleotide sequence;

(c) a second nucleotide sequence encoding a fusion protein comprising a first protein sequence containing at least one hepatitis B virus surface antigen epitope fused via a peptide bond to a second protein sequence containing at least one hepatitis B virus core antigen epitope; and

(d) a second promoter sequence operably linked to said second nucleotide sequence;

whereby the first and second nucleotide sequences are expressed by the virus in a suitable host.

5. The virus of claim 4 in which the fusion protein is selected from the group consisting of the following hepatitis B virus sequences: a preS1 region fused to a core antigen, a preS1 region fused to a core antigen derivative, a core antigen fused to S, a core antigen fused to a preS2 region, a core antigen derivative fused to S, a core antigen

derivative fused to a preS2 region, a core antigen fused to a preS1 region, a core antigen derivative fused to a preS1 region, and a core antigen fused to a preS1 region derivative; in which the core antigen derivative displays the antigenicity of the hepatitis B virus core antigen, and the preS1 region derivative displays the antigenicity of the hepatitis B virus preS1 region.

6. The virus of claim 4 in which the fusion protein is selected from the group consisting of a core-preS1 fragment fusion protein having the amino acid sequence shown in Figure 44 (SEQ ID NO:26), a core-preS2 fusion protein having the amino acid sequence shown in Figure 46 (SEQ ID NO:30), a preS1-coreΔ8 fusion protein having the amino acid sequence shown in Figure 76 (SEQ ID NO:45), and a preS1-core fusion protein having the amino acid sequence shown in Figure 78 (SEQ ID NO:49).

7. The virus of claim 6 in which the first nucleotide sequence comprises the sequence encoding MS as shown in Figure 23 (SEQ ID NO:7).

8. The virus of claim 6 in which the first nucleotide sequence encodes said derivative, said derivative being a fusion protein comprising a third protein sequence containing at least one hepatitis B virus surface antigen epitope, which third protein sequence is fused via a peptide bond to a fourth protein sequence containing at least one hepatitis B virus core antigen epitope; with the proviso that the first nucleotide sequence and the second nucleotide sequence have sequences that do not overlap by more than 20 nucleotides.

9. A recombinant replicable vaccinia virus, the genome of which comprises:

(a) a first nucleotide sequence encoding a hepatitis B virus MS antigen;

(b) a first promoter sequence operably linked to said first nucleotide sequence;

(c) a second nucleotide sequence encoding a fusion protein comprising a hepatitis B virus core antigen or a derivative thereof displaying core antigenicity, fused at its carboxy terminus via a peptide bond to the amino terminus of a hepatitis B virus preS1 region; and

(d) a second promoter sequence operably linked to said second nucleotide sequence;

whereby the first and second nucleotide sequences are expressed by the virus in a suitable host.

10. A recombinant replicable vaccinia virus, the genome of which comprises:

(a) a first nucleotide sequence encoding a hepatitis B virus MS antigen;

(b) a first promoter sequence operably linked to said first nucleotide sequence;

(c) a second nucleotide sequence encoding a fusion protein comprising a hepatitis B virus preS1 region fused at its carboxy terminus via a peptide bond to the amino terminus of a hepatitis B virus core antigen or a derivative thereof displaying core antigenicity; and

(d) a second promoter sequence operably linked to said second nucleotide sequence;

whereby the first and second nucleotide sequences are expressed by the virus in a suitable host.

11. The virus of claim 9 in which the core antigen or derivative thereof is a core antigen derivative having the amino acid sequence shown in Figure 33 (SEQ ID NO:15) or in Figure 30 (SEQ ID NO:13).

12. The virus of claim 10 in which the core antigen or derivative thereof is a core antigen derivative having the amino acid sequence shown in Figure 33 (SEQ ID NO:15) or in Figure 30 (SEQ ID NO:13).

13. The virus of claim 12 in which the first promoter is modified p7.5, and the second promoter is p7.5.

14. The virus of claim 12 in which the first promoter is p7.5, and the second promoter is p11.

15. A recombinant replicable vaccinia virus, the genome of which comprises:

(a) a first nucleotide sequence encoding a first fusion protein comprising a first hepatitis B virus core antigen derivative displaying core antigenicity fused at its carboxy terminus via a peptide bond to the amino terminus of a hepatitis B virus preS2 region;

(b) a first promoter sequence operably linked to said first nucleotide sequence;

(c) a second nucleotide sequence encoding a second fusion protein comprising a hepatitis B virus preS1 region fused at its carboxy terminus via a peptide bond to the amino terminus of a second hepatitis B virus core antigen derivative displaying core antigenicity; and

(d) a second promoter sequence operably linked to said second nucleotide sequence;

with the proviso that the first and second nucleotide sequences have sequences that do not overlap by more than 20 nucleotides; whereby the first and second nucleotide sequences are expressed by the virus in a suitable host.

16. The virus of claim 15 in which the core antigen or derivative thereof is a core antigen derivative having the amino acid sequence shown in Figure 33 (SEQ ID NO:15) or in Figure 30 (SEQ ID NO:13).

17. The virus of claim 16 in which the first promoter is modified p7.5, and the second promoter is p7.5.

18. A recombinant replicable vaccinia virus, the genome of which comprises a nucleotide sequence encoding a hepatitis B virus core antigen derivative having the amino acid sequence shown in Figure 30 (SEQ ID NO:13); and a promoter sequence

operably linked to said nucleotide sequence, whereby the nucleotide sequence is expressed by the virus in a suitable host.

19. A recombinant replicable vaccinia virus, the genome of which comprises a nucleotide sequence encoding a fusion protein, which protein comprises a first protein sequence containing at least one hepatitis B virus core antigen epitope, wherein said first protein sequence is fused at its carboxy terminus via a peptide bond to a second protein sequence containing at least one hepatitis B virus surface antigen epitope; and a promoter sequence operably linked to said nucleotide sequence, whereby the nucleotide sequence is expressed by the virus in a suitable host.

20. A recombinant replicable vaccinia virus, the genome of which comprises a nucleotide sequence encoding a fusion protein, which protein comprises a first protein sequence containing at least one hepatitis B virus surface antigen epitope, wherein said first protein sequence is fused at its carboxy terminus via a peptide bond to a second protein sequence containing at least one hepatitis B virus core antigen epitope; and a promoter sequence operably linked to said nucleotide sequence, whereby the nucleotide sequence is expressed by the virus in a suitable host.

21. The virus of claim 19 in which the fusion protein is selected from the group consisting of the following hepatitis B virus sequences: a core antigen or core antigen derivative fused to a preS1 region or preS1 derivative, core antigen or core

antigen derivative fused to a preS2 region or preS2 derivative, or a core antigen or core antigen derivative fused to an S protein; in which the core antigen derivative displays the antigenicity of the core antigen, the preS1 derivative displays the antigenicity of the preS1 region, the preS2 derivative displays the antigenicity of the preS2 region, and the S derivative displays the antigenicity of the S protein.

22. The virus of claim 1 in which the first and second nucleotide sequences are situated internally with respect to vaccinia virus thymidine kinase gene sequences, and the virus does not express a functional thymidine kinase.

23. The virus of claim 4 in which the first and second nucleotide sequences are situated internally with respect to vaccinia virus thymidine kinase gene sequences, and the virus does not express a functional thymidine kinase.

24. The virus of claim 9 in which the first and second nucleotide sequences are situated internally with respect to vaccinia virus thymidine kinase gene sequences, and the virus does not express a functional thymidine kinase.

25. The virus of claim 10 in which the first and second nucleotide sequences are situated internally with respect to vaccinia virus thymidine kinase gene sequences, and the virus does not express a functional thymidine kinase.

26. The virus of claim 19 in which the nucleotide sequence is situated internally with respect to vaccinia virus thymidine kinase gene sequences, and the virus does not express a functional thymidine kinase.

27. Plasmid pHTL-31, as deposited with the ATCC and assigned accession number 69124.

28. A recombinant replicable vaccinia virus, the genome of which comprises:

(a) a first nucleotide sequence that comprises the p7.5 promoter operably linked to a first coding sequence encoding a fusion protein consisting of a hepatitis B virus preS1 region fused to a hepatitis B virus core antigen derivative, as contained in plasmid pHTL-31; and

(b) a second nucleotide sequence that comprises (i) the modified p7.5 promoter operably linked to a second coding sequence encoding a hepatitis B virus MS antigen, and (ii) a sequence upstream of the initiator ATG of the MS antigen that is associated with strong ribosome binding, as contained in plasmid pHTL-31;

wherein plasmid pHTL-31 is as deposited with the ATCC and assigned accession no. 69124.

29. The virus of claim 1 which further comprises a sequence upstream of the initiator ATG in the first nucleotide sequence, that matches a

consensus sequence associated with strong ribosome binding.

30. The virus of claim 9 which further comprises a sequence upstream of the initiator ATG for the MS antigen in the first nucleotide sequence, that matches a consensus sequence associated with strong ribosome binding.

31. The virus of claim 10 which further comprises a sequence upstream of the initiator ATG for the MS antigen in the first nucleotide sequence, that matches a consensus sequence associated with strong ribosome binding.

32. A vaccine formulation comprising the recombinant vaccinia virus of claim 1; and a pharmaceutically acceptable carrier.

33. A vaccine formulation comprising the recombinant vaccinia virus of claim 4; and a pharmaceutically acceptable carrier.

34. A vaccine formulation comprising the recombinant vaccinia virus of claim 9; and a pharmaceutically acceptable carrier.

35. A vaccine formulation comprising the recombinant vaccinia virus of claim 10; and a pharmaceutically acceptable carrier.

36. A vaccine formulation comprising the recombinant vaccinia virus of claim 15; and a pharmaceutically acceptable carrier.

37. A vaccine formulation comprising a first recombinant virus of claim 1; a second recombinant virus of claim 1; and a pharmaceutically acceptable carrier; in which the first and second recombinant viruses encode different hepatitis B virus amino acid sequences.

38. A vaccine formulation comprising a first recombinant virus of claim 4; a second recombinant virus of claim 4; and a pharmaceutically acceptable carrier; in which the first and second recombinant viruses encode different hepatitis B virus amino acid sequences.

39. A vaccine formulation comprising a first recombinant virus of claim 19; a second recombinant virus of claim 19; and a pharmaceutically acceptable carrier; in which the first and second recombinant viruses express different fusion proteins.

40. A vaccine formulation comprising:

(a) a first recombinant replicable vaccinia virus, the genome of which comprises (i) a first nucleotide sequence encoding a first fusion protein, said first fusion protein comprising a hepatitis B virus core antigen or derivative thereof displaying core antigenicity, fused at its carboxy terminus via a peptide bind to a hepatitis B virus preS1 region or derivative thereof displaying preS1 antigenicity, and (ii) a first promoter sequence operably linked to said first nucleotide sequence, whereby the first

nucleotide sequence is expressed by said first virus in a suitable host;

(b) a second recombinant replicable vaccinia virus, the genome of which comprises (i) a second nucleotide sequence encoding a second fusion protein, said second fusion protein comprising a hepatitis B virus core antigen or derivative thereof displaying core antigenicity, fused at its carboxy terminus via a peptide bind to a hepatitis B virus preS2 region or derivative thereof displaying preS2 antigenicity, and (ii) a second promoter sequence operably linked to said second nucleotide sequence, whereby the second nucleotide sequence is expressed by said second virus in a suitable host;

(c) a third recombinant replicable vaccinia virus, the genome of which comprises (i) a third nucleotide sequence encoding a hepatitis B virus S protein or a derivative thereof displaying S antigenicity, and (ii) a third promoter sequence operably linked to said third nucleotide sequence, whereby the third nucleotide sequence is expressed by said third virus in a suitable host; and

(d) a pharmaceutically acceptable carrier.

41. The vaccine formulation of claim 40 in which the first fusion protein has the amino acid sequence (SEQ ID NO:26) depicted in Figure 44, the

first promoter is modified p7.5, the second fusion protein has the amino acid sequence (SEQ ID NO:30) depicted in Figure 46, the second promoter is modified p7.5, the third nucleotide sequence encodes a protein having the amino acid sequence (SEQ ID NO:10) depicted in Figure 24, and the third promoter is modified p7.5.

42. A vaccine formulation comprising:

(a) a first recombinant replicable vaccinia virus, the genome of which comprises (i) a first nucleotide sequence encoding a first fusion protein, said first fusion protein comprising a hepatitis B virus core antigen or derivative thereof displaying core antigenicity, fused at its carboxy terminus via a peptide bind to a hepatitis B virus preS1 region or derivative thereof displaying preS1 antigenicity, and (ii) a first promoter sequence operably linked to said first nucleotide sequence, whereby the first nucleotide sequence is expressed by said first virus in a suitable host;

(b) a second recombinant replicable vaccinia virus, the genome of which comprises (i) a second nucleotide sequence encoding a hepatitis B virus MS protein or a derivative thereof displaying MS antigenicity, and (ii) a second promoter operably linked to said second nucleotide sequence, whereby the second nucleotide sequence is expressed by said second virus in a suitable host; and

(c) a pharmaceutically acceptable carrier.

43. The vaccine formulation of claim 42 in which the first fusion protein has the amino acid sequence (SEQ ID NO:26) depicted in Figure 44, the first promoter is modified p7.5, the second nucleotide sequence encodes an MS protein having the sequence (SEQ ID NO:8) depicted in Figure 23, and in which a sequence upstream of the second nucleotide sequence matches a consensus sequence associated with strong ribosome binding.

44. A vaccine formulation comprising:

(a) a first recombinant replicable vaccinia virus, the genome of which comprises (i) a first nucleotide sequence encoding a first fusion protein, said first fusion protein comprising a hepatitis B virus core antigen or derivative thereof displaying core antigenicity, fused at its carboxy terminus via a peptide bind to a hepatitis B virus preS2 region or derivative thereof displaying preS2 antigenicity, and (ii) a first promoter sequence operably linked to said first nucleotide sequence, whereby the first nucleotide sequence is expressed by said first virus in a suitable host;

(b) a second recombinant replicable vaccinia virus, the genome of which comprises (i) a second nucleotide sequence encoding a hepatitis B virus S protein or a derivative thereof displaying S antigenicity, and (ii) a

second promoter sequence operably linked to said second nucleotide sequence, whereby the second nucleotide sequence is expressed by said second virus in a suitable host; and

(c) a pharmaceutically acceptable carrier.

45. The vaccine formulation of claim 44 in which the first fusion protein has the amino acid sequence (SEQ ID NO:30) depicted in Figure 46, the first promoter is modified p7.5, the second nucleotide sequence encodes a protein having the amino acid sequence (SEQ ID NO:10) depicted in Figure 24, and the second promoter is modified p7.5.

46. A method of preventing hepatitis or other undesirable consequence of hepatitis B virus infection in a subject comprising administering to the subject an effective amount of the recombinant virus of claim 1.

47. A method of preventing hepatitis or other undesirable consequence of hepatitis B virus infection in a subject comprising administering to the subject an effective amount of the recombinant virus of claim 4.

48. A method of preventing hepatitis or other undesirable consequence of hepatitis B virus infection in a subject comprising administering to the subject an effective amount of the recombinant virus of claim 9.

49. A method of preventing hepatitis or other undesirable consequence of hepatitis B virus

infection in a subject comprising administering to the subject an effective amount of the recombinant virus of claim 10.

50. A method of preventing hepatitis or other undesirable consequence of hepatitis B virus infection in a subject comprising administering to the subject an effective amount of the recombinant virus of claim 15.

51. A method of preventing hepatitis or other undesirable consequence of hepatitis B virus infection in a subject comprising administering to the subject an effective amount of the vaccine formulation of claim 39.

52. A method of preventing hepatitis or other undesirable consequence of hepatitis B virus infection in a subject comprising administering to the subject an effective amount of the vaccine formulation of claim 40.

53. A method of preventing hepatitis or other undesirable consequence of hepatitis B virus infection in a subject comprising administering to the subject an effective amount of:

(a) a first recombinant replicable vaccinia virus, the genome of which comprises (i) a first nucleotide sequence encoding a first fusion protein, said first fusion protein comprising a hepatitis B virus core antigen or derivative thereof displaying core antigenicity, fused at its carboxy terminus via a peptide bind to a

hepatitis B virus preS1 region or derivative thereof displaying preS1 antigenicity, and (ii) a first promoter sequence operably linked to said first nucleotide sequence, whereby the first nucleotide sequence is expressed by said first virus in a suitable host;

(b) a second recombinant replicable vaccinia virus, the genome of which comprises (i) a second nucleotide sequence encoding a second fusion protein, said second fusion protein comprising a hepatitis B virus core antigen or derivative thereof displaying core antigenicity, fused at its carboxy terminus via a peptide bind to a hepatitis B virus preS2 region or derivative thereof displaying preS2 antigenicity, and (ii) a second promoter sequence operably linked to said second nucleotide sequence, whereby the second nucleotide sequence is expressed by said second virus in a suitable host; and

(c) a third recombinant replicable vaccinia virus, the genome of which comprises (i) a third nucleotide sequence encoding a hepatitis B virus S protein or a derivative thereof displaying S antigenicity, and (ii) a third promoter sequence operably linked to said third nucleotide sequence, whereby the third nucleotide sequence is expressed by said third virus in a suitable host.

54. A method of preventing hepatitis or other undesirable consequence of hepatitis B virus infection in a subject comprising administering to the subject an effective amount of:

(a) a first recombinant replicable vaccinia virus, the genome of which comprises (i) a first nucleotide sequence encoding a first fusion protein, said first fusion protein comprising a hepatitis B virus core antigen or derivative thereof displaying core antigenicity, fused at its carboxy terminus via a peptide bind to a hepatitis B virus preS1 region or derivative thereof displaying preS1 antigenicity, and (ii) a first promoter sequence operably linked to said first nucleotide sequence, whereby the first nucleotide sequence is expressed by said first virus in a suitable host; and

(b) a second recombinant replicable vaccinia virus, the genome of which comprises (i) a second nucleotide sequence encoding a hepatitis B virus MS protein or a derivative thereof displaying MS antigenicity, and (ii) a second promoter operably linked to said second nucleotide sequence, whereby the second nucleotide sequence is expressed by said second virus in a suitable host.

55. A method of preventing hepatitis or other undesirable consequence of hepatitis B virus

infection in a subject comprising administering to the subject an effective amount of:

(a) a first recombinant replicable vaccinia virus, the genome of which comprises (i) a first nucleotide sequence encoding a first fusion protein, said first fusion protein comprising a hepatitis B virus core antigen or derivative thereof displaying core antigenicity, fused at its carboxy terminus via a peptide bind to a hepatitis B virus preS2 region or derivative thereof displaying preS2 antigenicity, and (ii) a first promoter sequence operably linked to said first nucleotide sequence, whereby the first nucleotide sequence is expressed by said first virus in a suitable host; and

(b) a second recombinant replicable vaccinia virus, the genome of which comprises (i) a second nucleotide sequence encoding a hepatitis B virus S protein or a derivative thereof displaying S antigenicity, and (ii) a second promoter sequence operably linked to said second nucleotide sequence, whereby the second nucleotide sequence is expressed by said second virus in a suitable host.

56. A method of treating hepatitis or other undesirable consequence of hepatitis B virus infection comprising administering to a subject in need of such

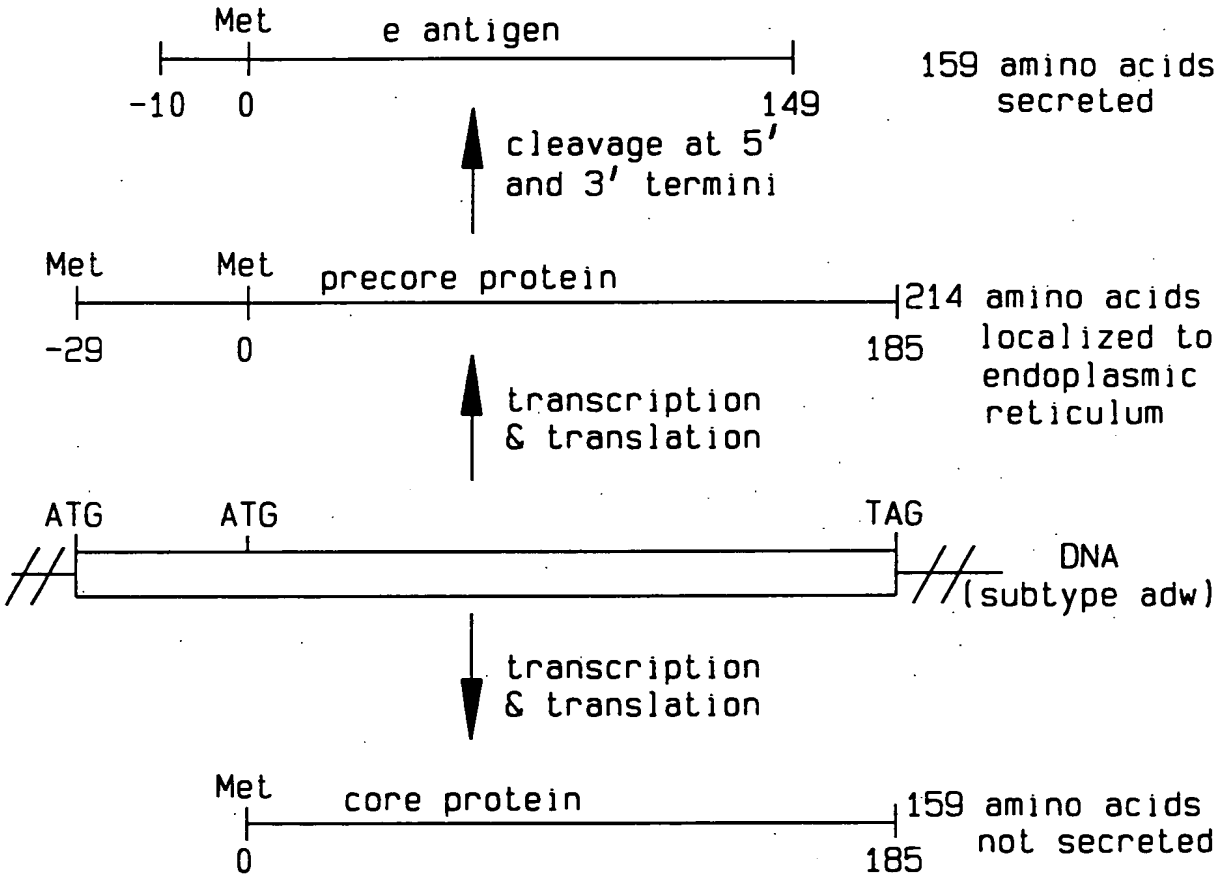
treatment an effective amount of the recombinant vaccinia virus of claim 1.

57. A method of treating hepatitis or other undesirable consequence of hepatitis B virus infection comprising administering to a subject in need of such treatment an effective amount of the recombinant vaccinia virus of claim 4.

58. A purified protein having the amino acid sequence (SEQ ID NO:13) depicted in Figure 30.

59. A purified nucleic acid encoding the protein of claim 58.

FIG. 1


Met
e antigen
-10
0
149
159 amino acids secreted
cleavage at 5′ and 3′ termini
Met
Met
precore protein
-29
0
185
214 amino acids localized to endoplasmic reticulum
transcription & translation
ATG
ATG
TAG
DNA (subtype adw)
transcription & translation
Met
core protein
0
185
159 amino acids not secreted

```
        |   10      |   20      |   30      |   40      |   50      |   60
  1  ATAAATAATA AATACAATAA TTAATTTCTC GTAAAAGTAG AAAATATATT CTAATTTATT  60
 61  GCAC                                                              64
        |   10      |   20      |   30      |   40      |   50      |   60
```

FIG.2

```
        |   10      |   20      |   30      |   40      |   50      |   60
  1  CTAGAAGCGA TGCTACGCTA GTCACAATCA CCACTTTCAT ATTTAGAATA TATGTATGTA  60
 61  AAAATATAGT AGAATTTCAT TTTGTTTTTT TCTATGCTAT AAAT                  104
        |   10      |   20      |   30      |   40      |   50      |   60
```

FIG.4

```
        |   10      |   20      |   30      |   40      |   50      |   60
  1  AAAAATTGAA ATTTTATTTT TTTTTTTTGG AATATAAATA AG                    42
        |   10      |   20      |   30      |   40      |   50      |   60
```

FIG.6

FIG. 3

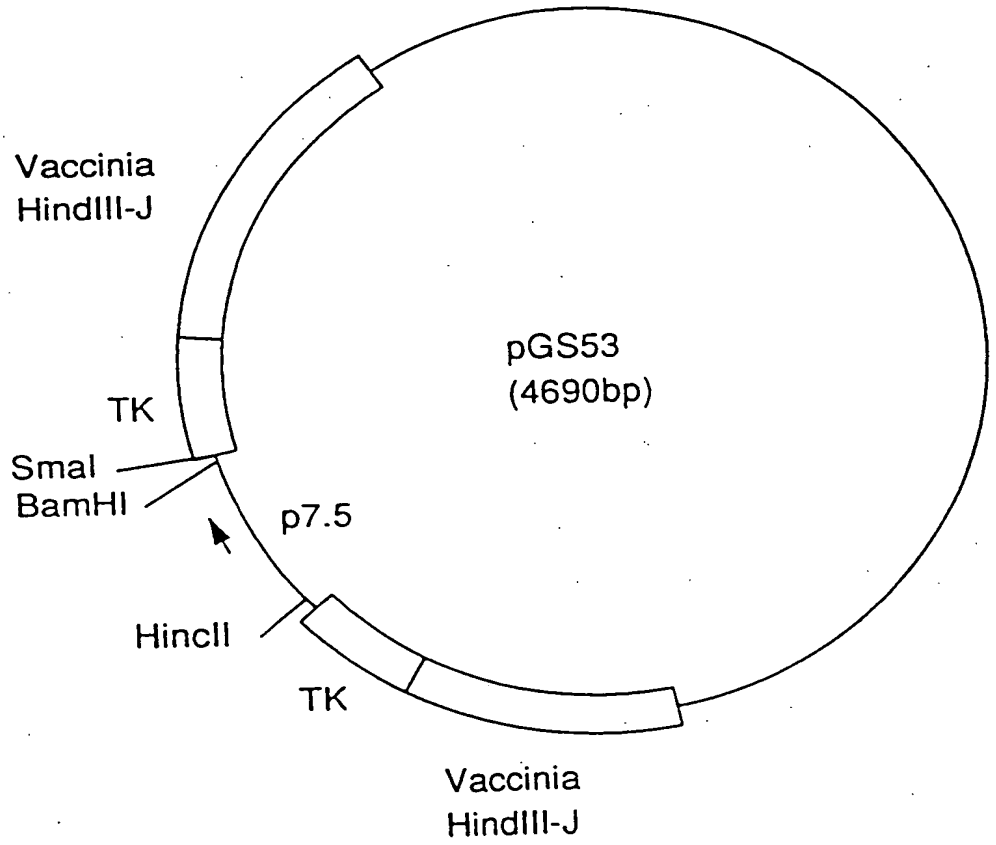

Vaccinia
HindIII-J
pGS53
(4690bp)
TK
SmaI
BamHI
p7.5
HincII
TK
Vaccinia
HindIII-J

FIG. 5

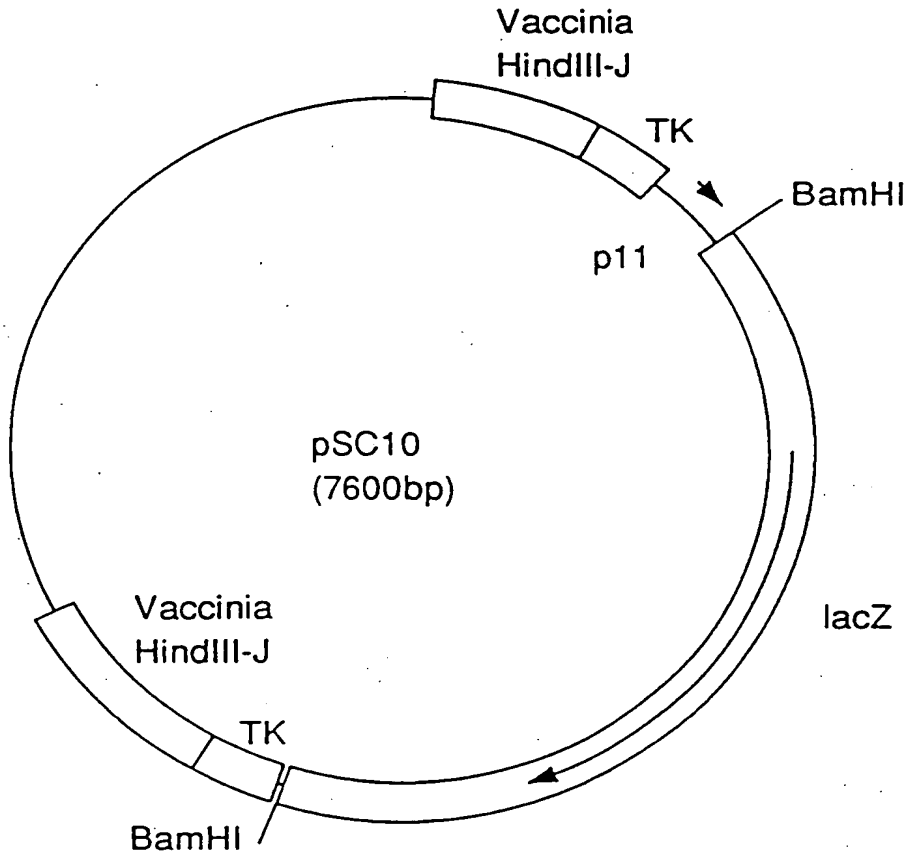

Vaccinia
HindIII-J
TK
BamHI
p11
pSC10
(7600bp)
lacZ
Vaccinia
HindIII-J
TK
BamHI

FIG. 7

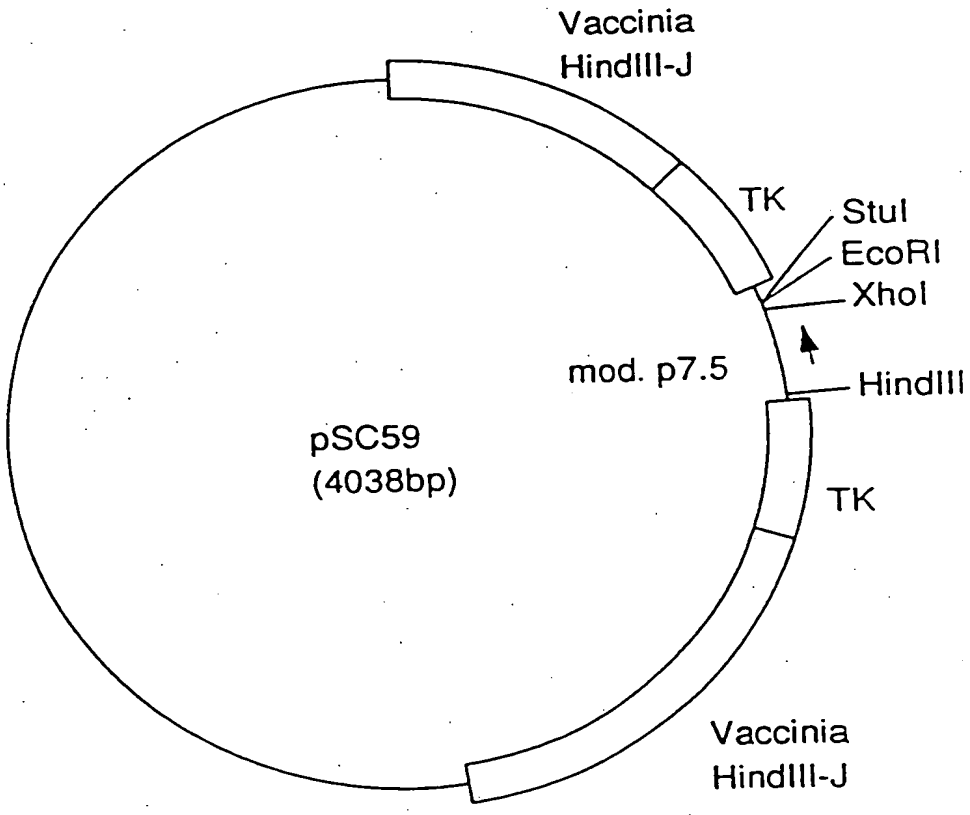

Vaccinia
HindIII-J
TK
StuI
EcoRI
XhoI
mod. p7.5
HindIII
pSC59
(4038bp)
TK
Vaccinia
HindIII-J

FIG. 8

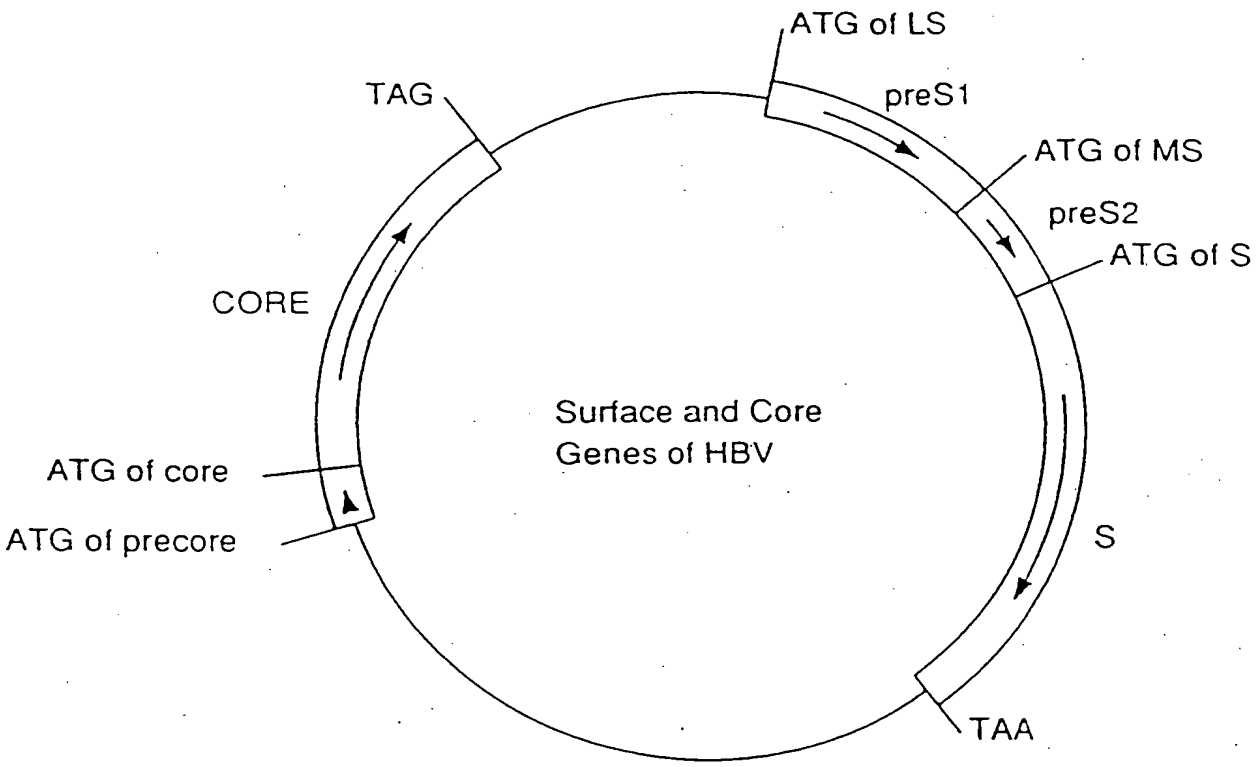
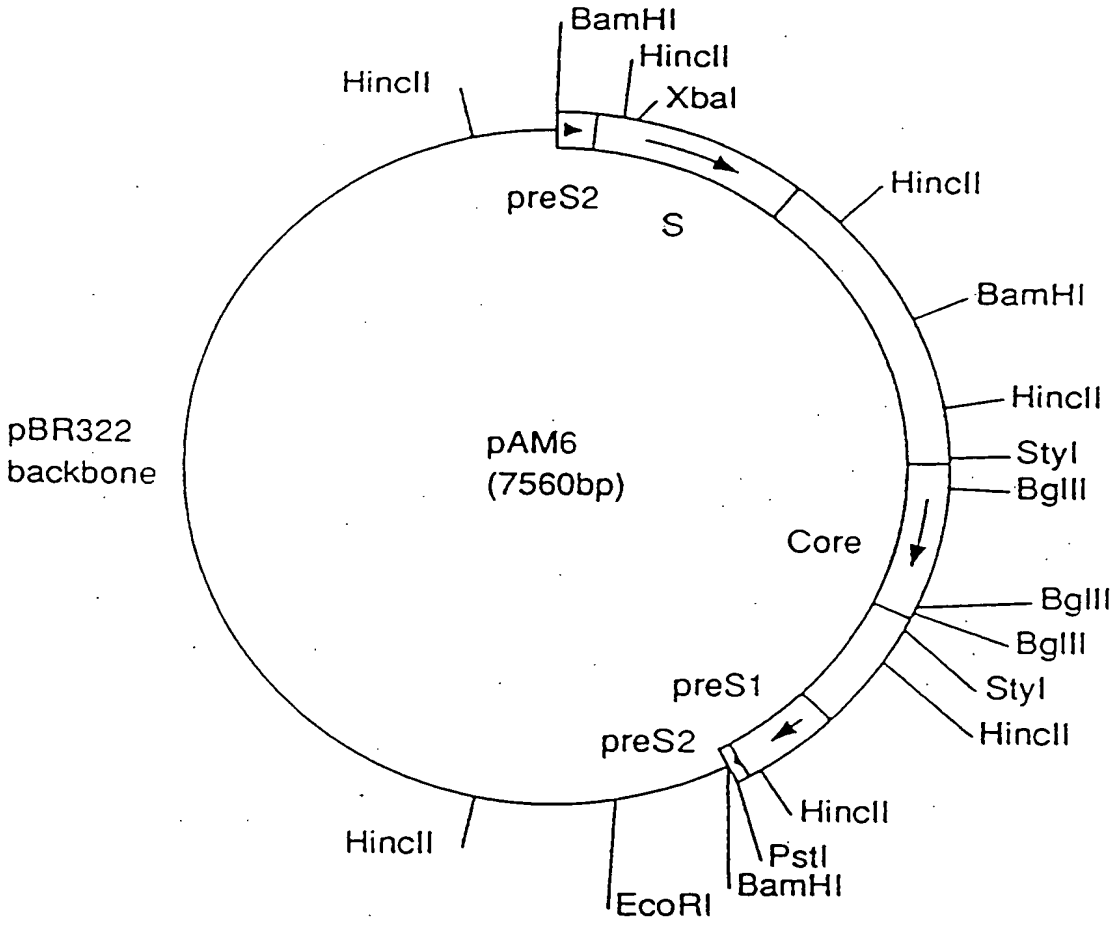

ATG of LS
TAG
preS1
ATG of MS
preS2
ATG of S
CORE
Surface and Core
Genes of HBV
ATG of core
ATG of precore
S
TAA

FIG. 9


BamHI
HincII
XbaI
HincII
preS2
S
HincII
BamHI
HincII
StyI
BglII
pBR322
backbone
pAM6
(7560bp)
Core
BglII
BglII
StyI
HincII
preS1
preS2
HincII
PstI
BamHI
HincII
EcoRI

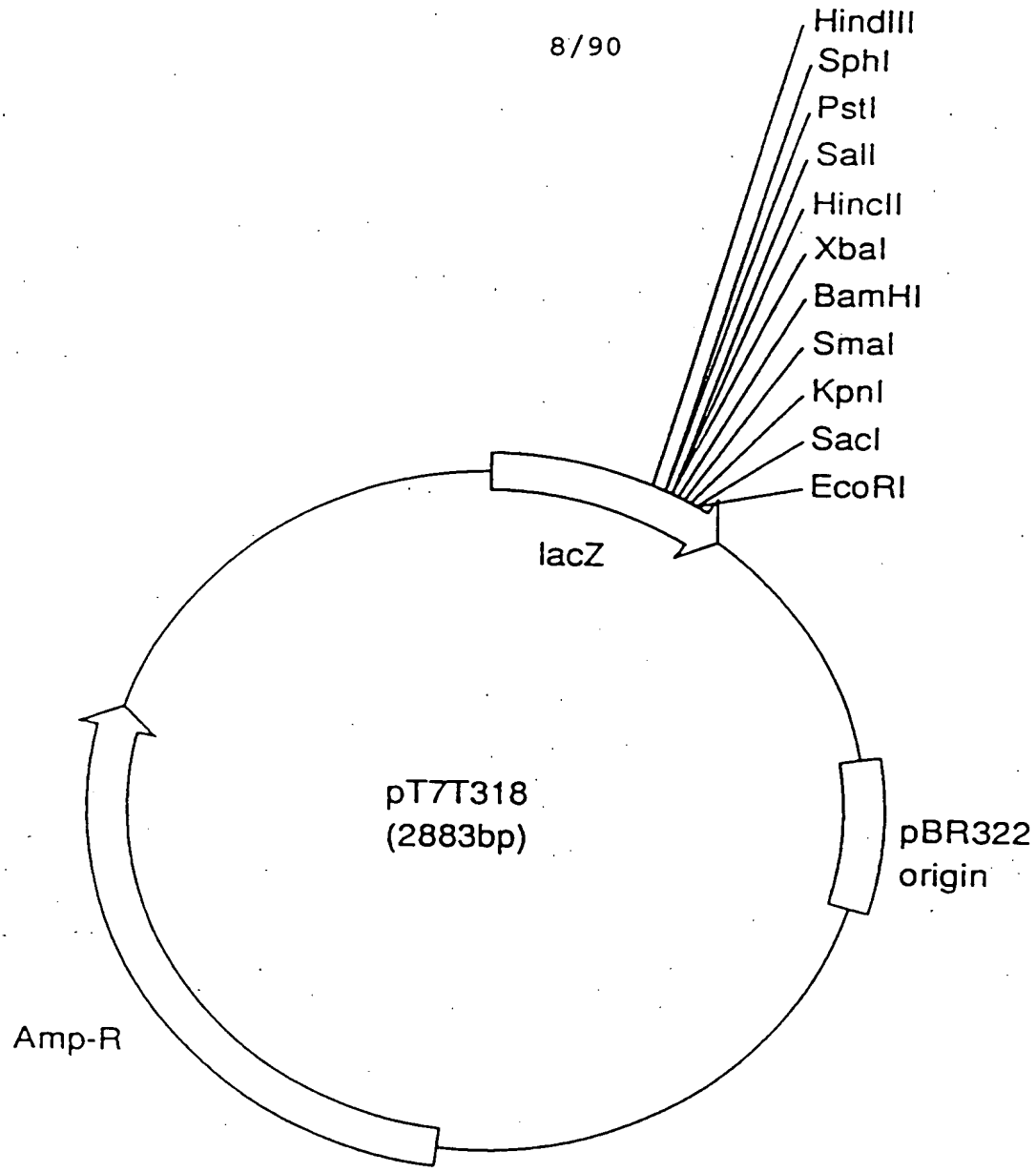
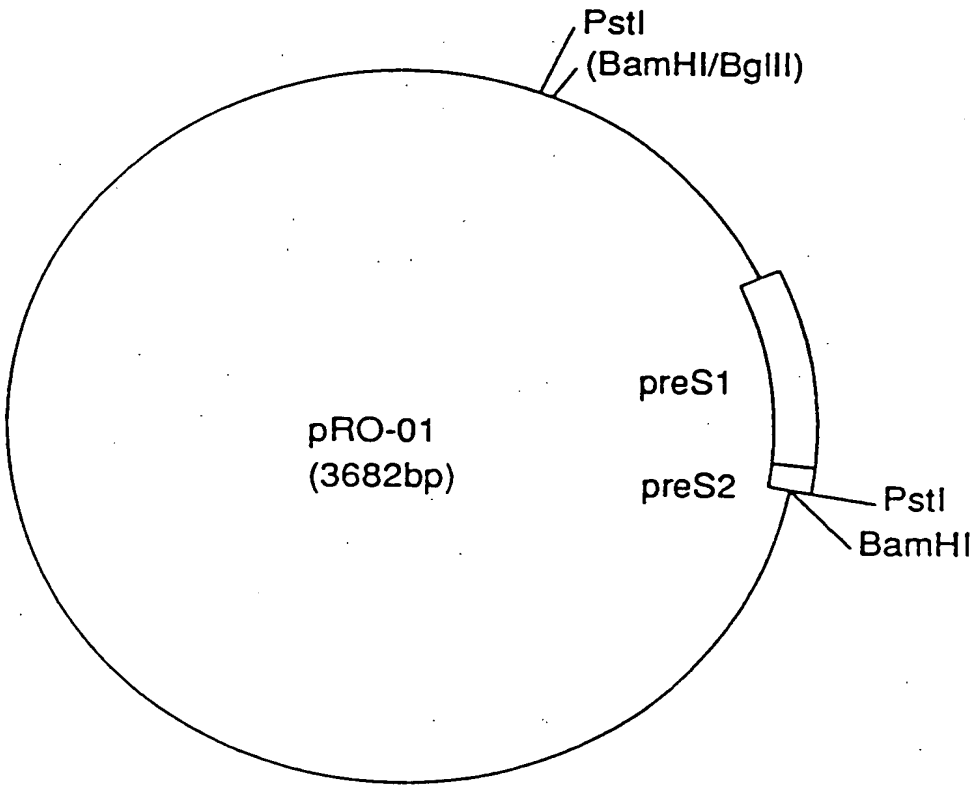

HindIII
SphI
PstI
SalI
HincII
XbaI
BamHI
SmaI
KpnI
SacI
EcoRI
lacZ
pT7T318
(2883bp)
pBR322
origin
Amp-R

FIG. 10

FIG. 11



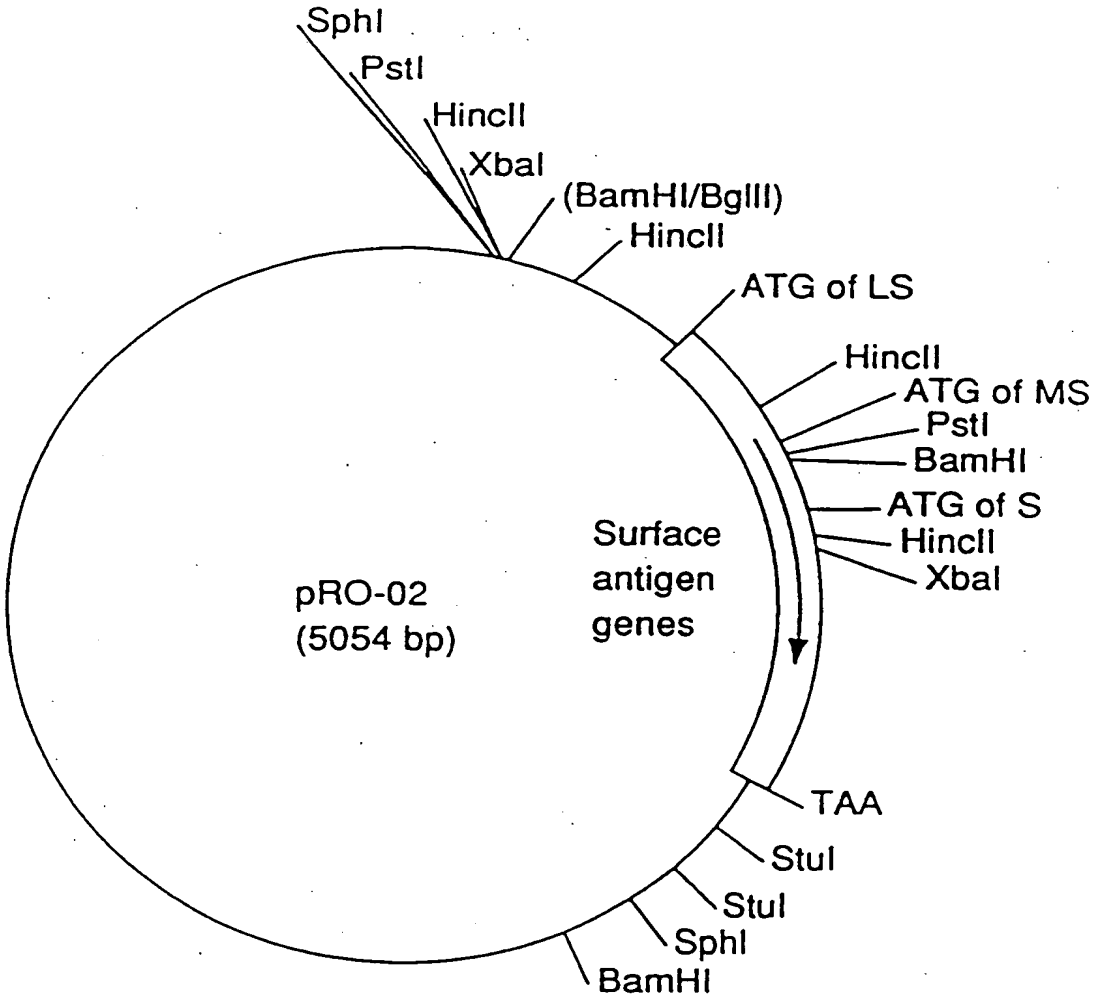
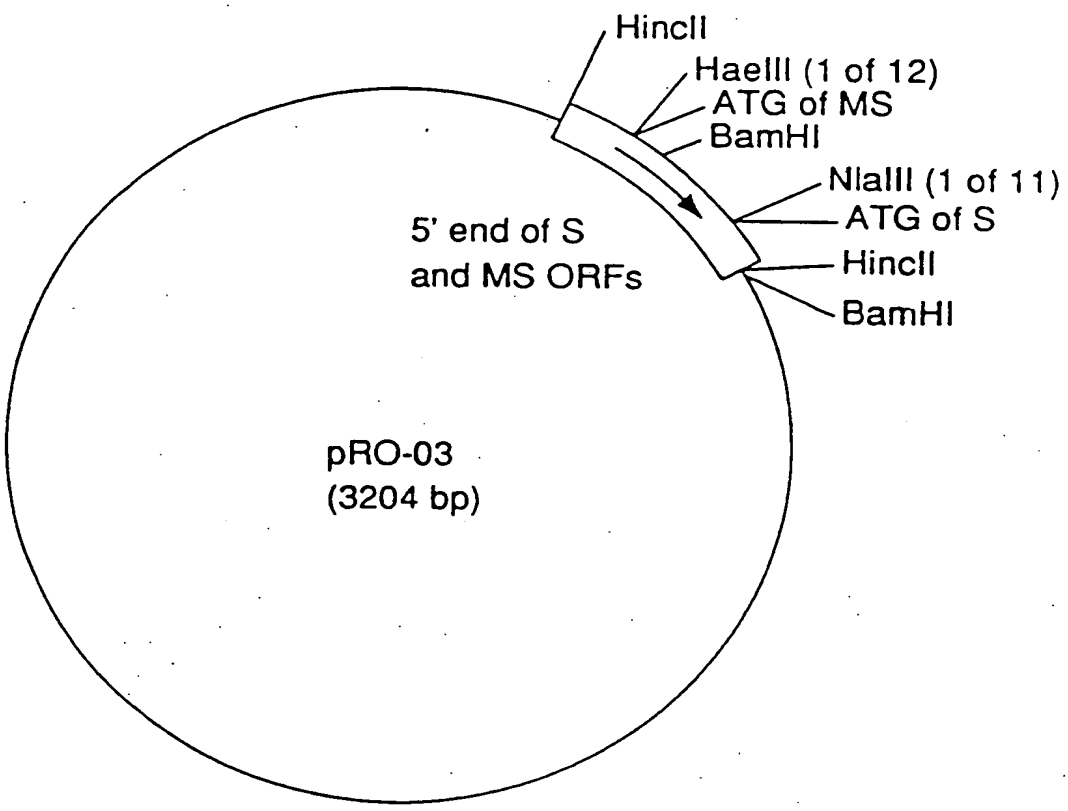

SphI
PstI
HincII
XbaI
(BamHI/BglII)
HincII
ATG of LS
HincII
ATG of MS
PstI
BamHI
ATG of S
HincII
XbaI
Surface
antigen
genes
pRO-02
(5054 bp)
TAA
StuI
StuI
SphI
BamHI

FIG. 12

FIG. 13


HincII
HaeIII (1 of 12)
ATG of MS
BamHI
NlaIII (1 of 11)
ATG of S
HincII
BamHI
5' end of S
and MS ORFs
pRO-03
(3204 bp)

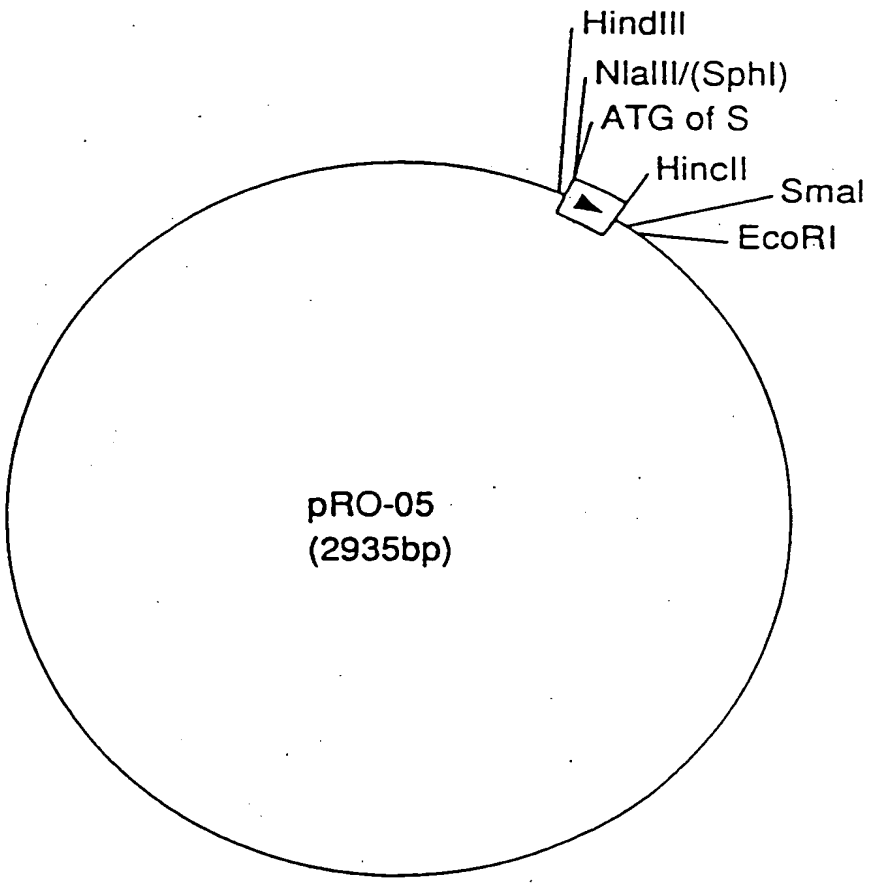


FIG. 14

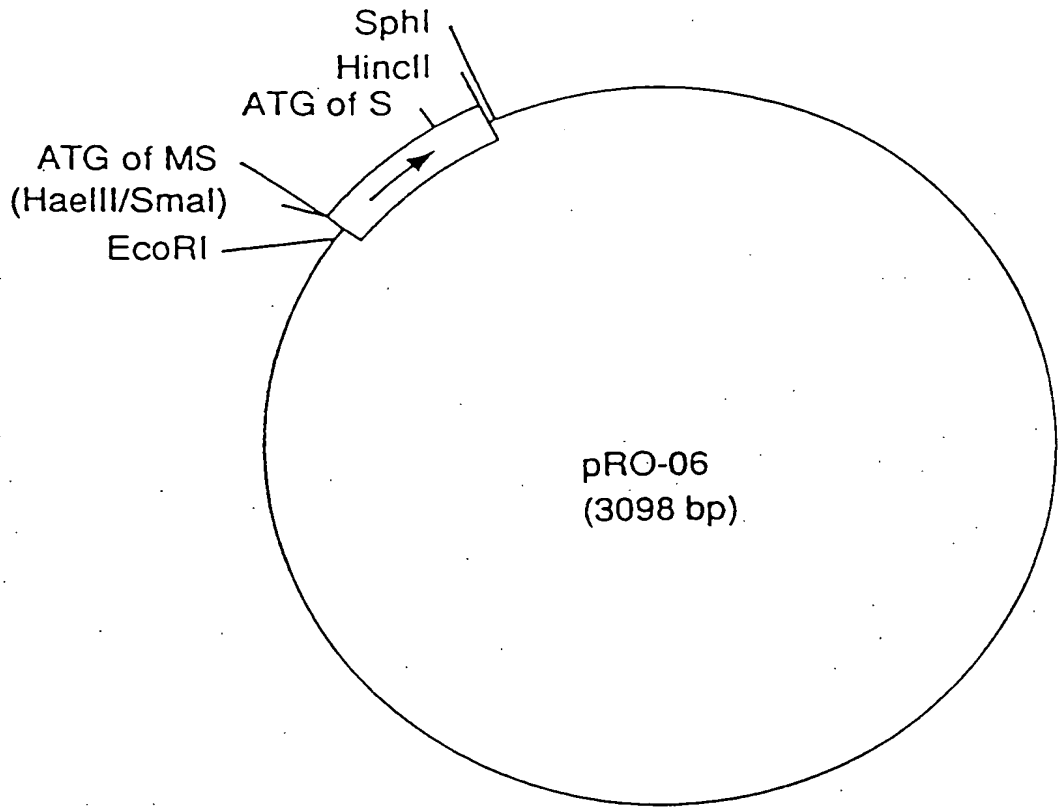


FIG. 15

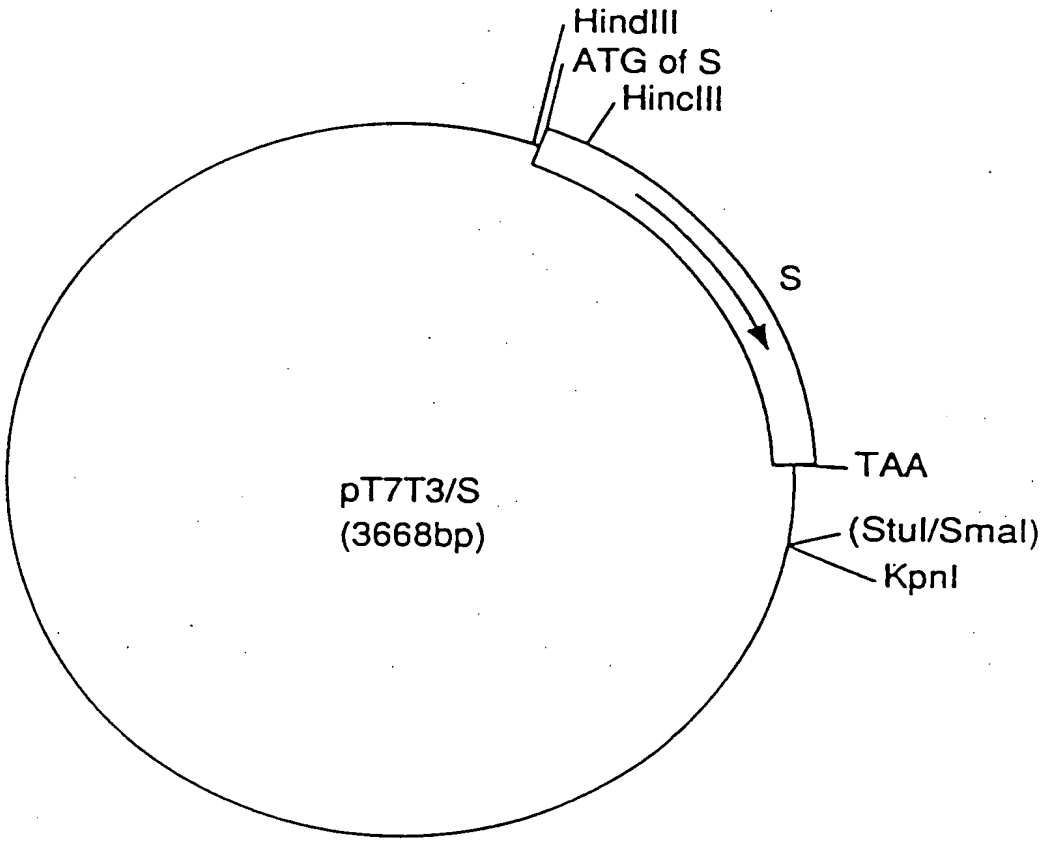


FIG. 16

FIG. 17

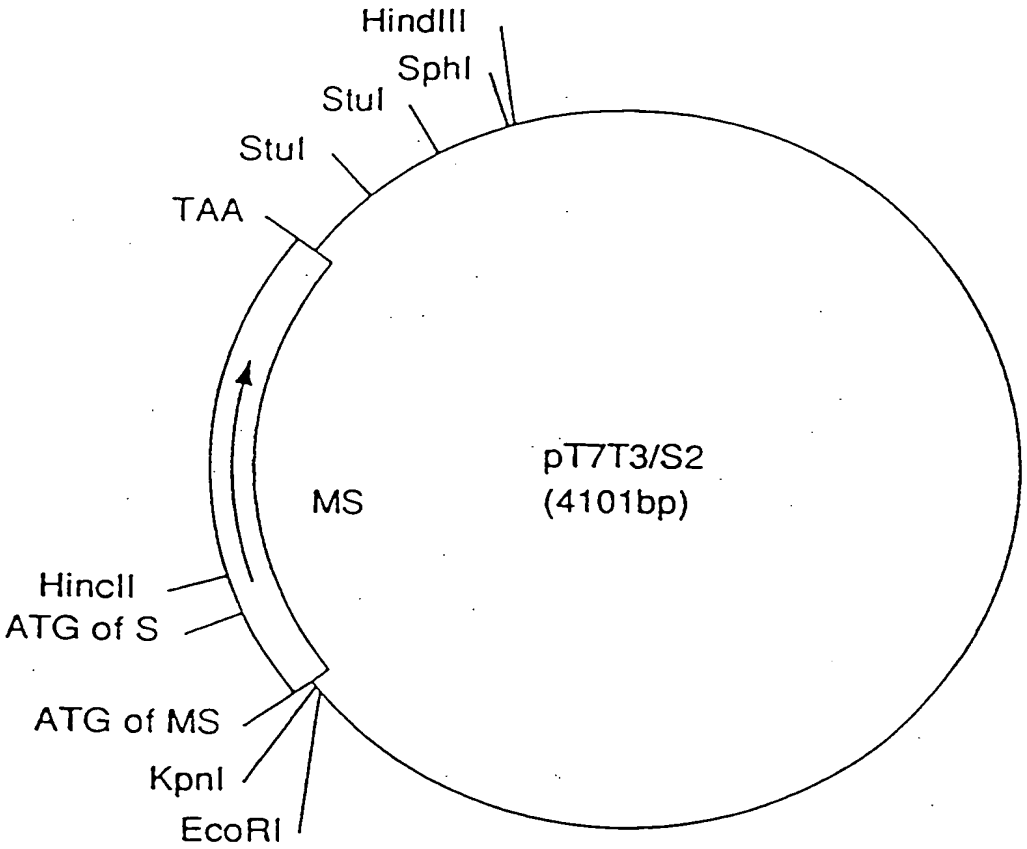

HindIII
SphI
StuI
StuI
TAA
MS
pT7T3/S2
(4101bp)
HincII
ATG of S
ATG of MS
KpnI
EcoRI

FIG. 18

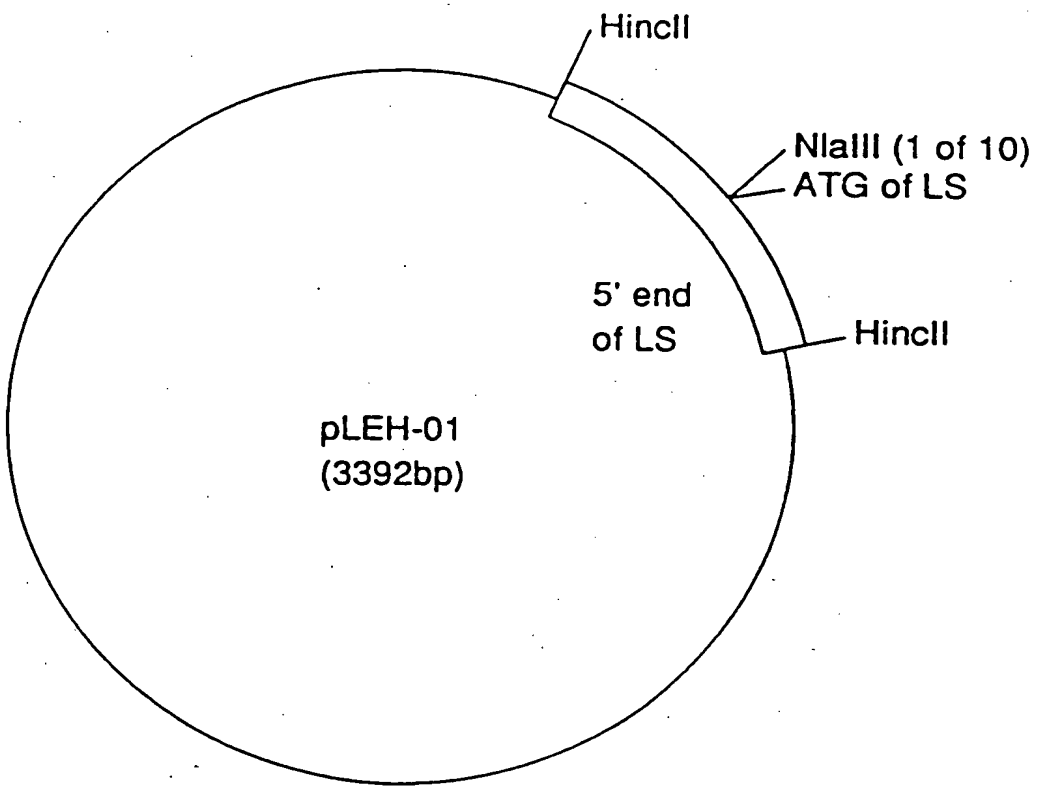

HincII
NlaIII (1 of 10)
ATG of LS
5' end
of LS
HincII
pLEH-01
(3392bp)

FIG. 19

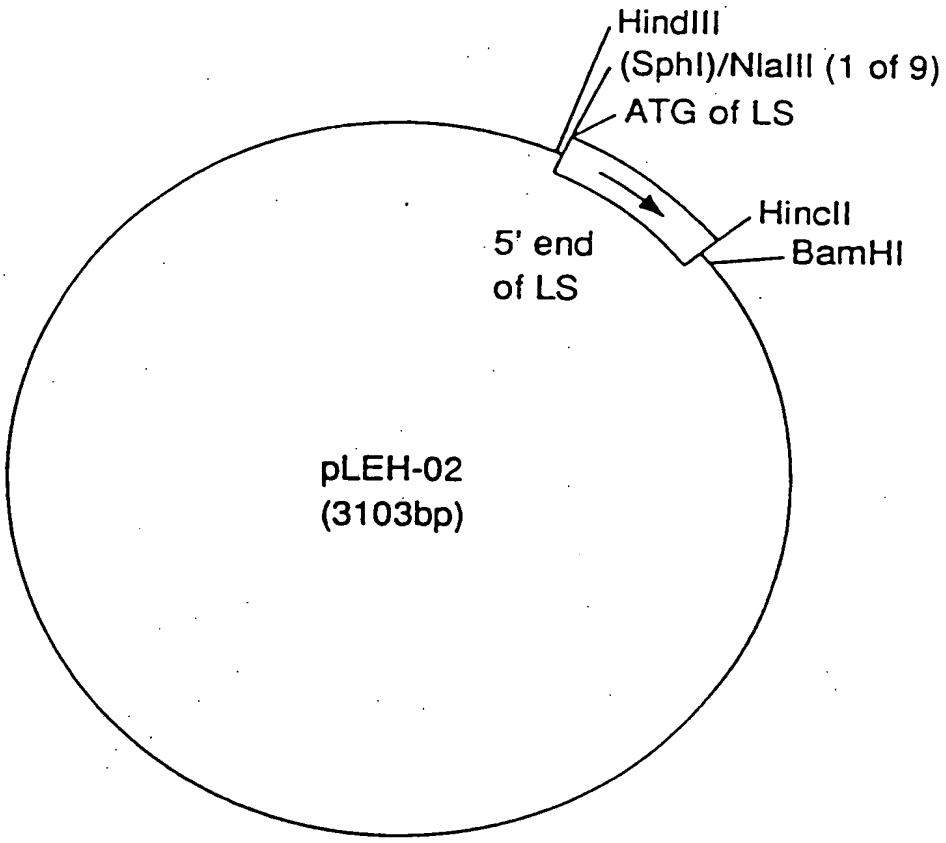

HindIII
(SphI)/NlaIII (1 of 9)
ATG of LS
HincII
BamHI
5' end
of LS
pLEH-02
(3103bp)

FIG. 20

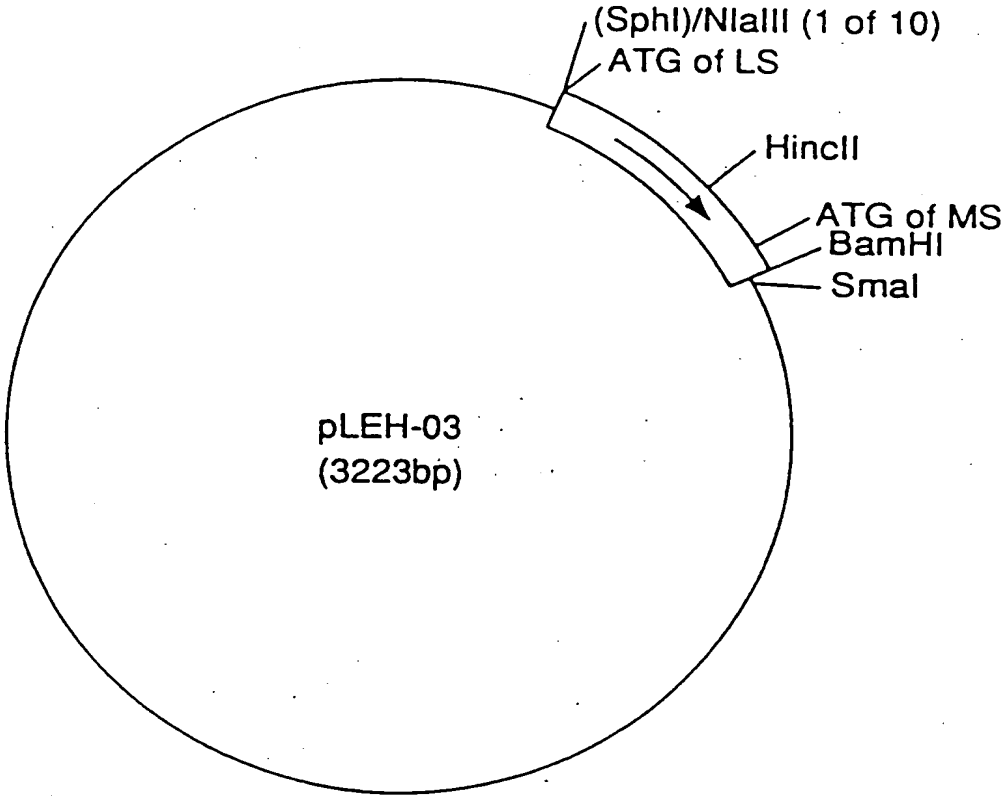
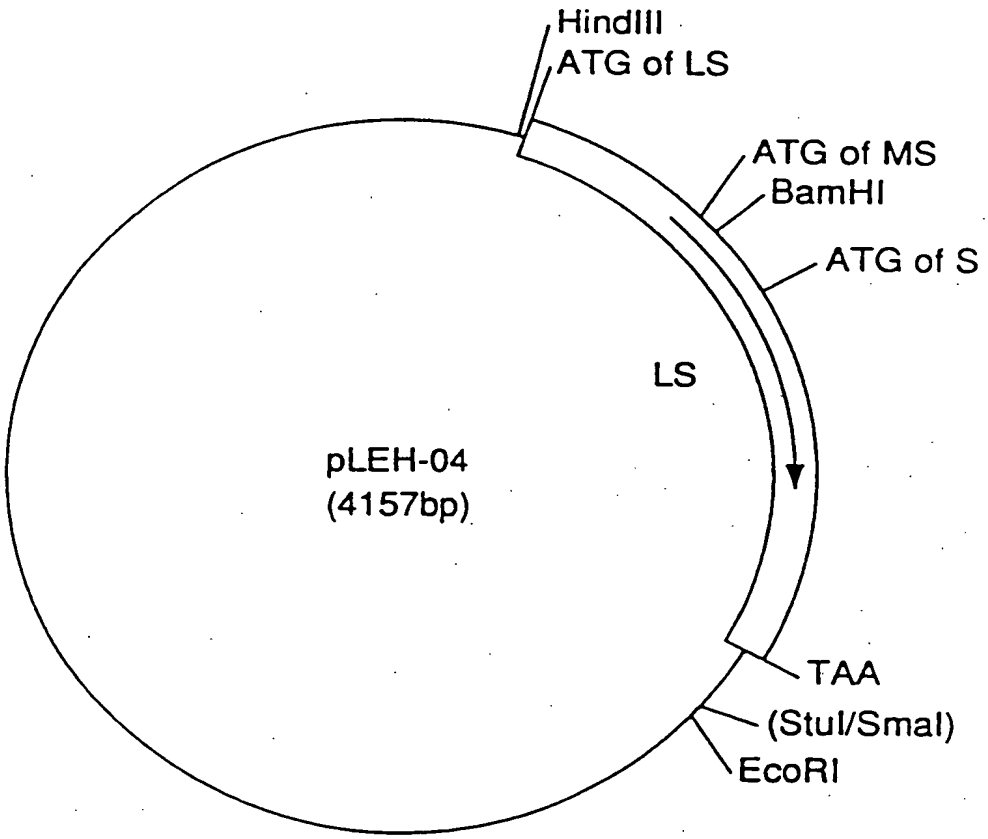

(SphI)/NlaIII (1 of 10)
ATG of LS
HincII
ATG of MS
BamHI
SmaI
pLEH-03
(3223bp)

FIG. 21


HindIII
ATG of LS
ATG of MS
BamHI
ATG of S
LS
pLEH-04
(4157bp)
TAA
(StuI/SmaI)
EcoRI

```
1   /   1                                31   /   11
ATG GGG ACG AAT CTT TCT GTT CCC AAt CCT CTG GGA TTC tTT CCC GAT CAT CAG TTG GAC
met gly thr asn leu ser val pro asn pro leu gly phe phe pro asp his gln leu asp
61   /   21                              91   /   31
CCT GCA TTC GGA GCC AAC TCA Aac AAT CCA GAT TGG GAC TTC AAC CCC ATC AAG GAC CAC
pro ala phe gly ala asn ser asn asn pro asp trp asp phe asn pro ile lys asp his
121   /   41                            151   /   51
TGG CCA GCA GCC AAC CAG GTA GGA GTG GGA GCA TTC GGG CCA GGG tTC ACC CCT CCA CAC
trp pro ala ala asn gln val gly val gly ala phe gly pro gly phe thr pro pro his
181   /   61                            211   /   71
GGC GGT gTT TTG GGG TGG AGC CCT CAG GCT CAG GGC ATA TTG ACC ACA GTG TCA ACA ATT
gly gly val leu gly trp ser pro gln ala gln gly ile leu thr thr val ser thr ile
241   /   81                            271   /   91
CCT CCT CCT GCC TCC ACC AAT CGG CAG TCA GGA AGG CAG CCT ACT CCC ATC TCT CCA CCT
pro pro pro ala ser thr asn arg gln ser gly arg gln pro thr pro ile ser pro pro
301   /   101                           331   /   111
CTA AGA GAC AGT CAT CCT CAG GCC ATG CAG TGG AAc TCC ACT GCC TTc CAC CAA GCT CTG
leu arg asp ser his pro gln ala met gln trp asn ser thr ala phe his gln ala leu
361   /   121                           391   /   131
CAG GAT CCC AGA GTC AGG GGT CTG TAT tTT CCT GCT GGT GGC TCC AGT TCA GGA ACA GTA
gln asp pro arg val arg gly leu tyr phe pro ala gly gly ser ser ser gly thr val
421   /   141                           451   /   151
AAC CCT GCT CCG AAT ATT GCC TCT CAC ATC TCG TCA ATC TCC GCG AGG ACT GGG GAC CCT
asn pro ala pro asn ile ala ser his ile ser ser ile ser ala arg thr gly asp pro
481   /   161                           511   /   171
GTG ACG AaC ATG GAG AAC ATC ACA TCA GGA TTC CTA GGA CCC CTG CTC GTG TTA CAG GCG
val thr asn met glu asn ile thr ser gly phe leu gly pro leu leu val leu gln ala
541   /   181                           571   /   191
GGG TTT TTC TTG TTG ACA AGA ATC CTC ACA ATA CCG CAG AGT CTA GAC TCG TGG TGG ACT
gly phe phe leu leu thr arg ile leu thr ile pro gln ser leu asp ser trp trp thr
601   /   201                           631   /   211
TCT CTC AAT TTT CTA GGG GGA TCA CCC GTG TGT CTT GGC CAA AAT TCG CAG TCC CCA ACC
ser leu asn phe leu gly gly ser pro val cys leu gly gln asn ser gln ser pro thr
661   /   221                           691   /   231
TCC AAT CAC TCA CCA ACC TCC TGT CCT CCA ATT TGT CCT GGT TAT CGC TGG ATG TGT CTG
ser asn his ser pro thr ser cys pro pro ile cys pro gly tyr arg trp met cys leu
721   /   241                           751   /   251
CGG CGT TTT ATC ATA TTC CTC TTC ATC CTG CTG CTA TGC CTC ATC TTC TTA TTG GTT CTT
arg arg phe ile ile phe leu phe ile leu leu leu cys leu ile phe leu leu val leu
```

FIG.22A

```
781  /  261                                  811  /  271
CTG GAT TAT CAA GGT ATG TTG CCC GTT TGT CCT CTA ATT CCA GGA TCA ACA ACA ACC AGT
leu asp tyr gln gly met leu pro val cys pro leu ile pro gly ser thr thr thr ser
841  /  281                                  871  /  291
ACG GGA CCA TGC AAA ACC TGC ACG ACT CCT GCT CAA GGC AAC TCT AtG TTT CCC TCA TGT
thr gly pro cys lys thr cys thr thr pro ala gln gly asn ser met phe pro ser cys
901  /  301                                  931  /  311
TGC TGT ACA AAA CCT ACG GAT GGA AAT TGC ACC TGT ATT CCC ATC CCA TCG TCt TGG GCT
cys cys thr lys pro thr asp gly asn cys thr cys ile pro ile pro ser ser trp ala
961  /  321                                  991  /  331
TTC GCA AAA TAC CTA TGG GAG TGG GCC TCA GTC CGT TTC TCT TGG CTC AGT TTA CTA GTG
phe ala lys tyr leu trp glu trp ala ser val arg phe ser trp leu ser leu leu val
1021 /  341                                  1051 /  351
CCA TTT GTT CAG TGG TTC GTA GGG CTT TCC CCC ACT GTT TGG CTT TCA GCT ATA TGG ATG
pro phe val gln trp phe val gly leu ser pro thr val trp leu ser ala ile trp met
1081 /  361                                  1111 /  371
ATG TGG TAT TGG GGG CCA AGT CTG TAC AGC ATC GTG AGT CCC TTT ATA CCG CTG TTA CCA
met trp tyr trp gly pro ser leu tyr ser ile val ser pro phe ile pro leu leu pro
1141 /  381
ATT TTC TTT TGT CTC TGG GTA TAC ATT TAA
ile phe phe cys leu trp val tyr ile OCH
```

FIG.22B

```
1   /   1                                   31  /  11
ATG CAG TGG AAC TCC ACT GCC TTC CAC CAA GCT CTG CAG GAT CCC AGA GTC AGG GGT CTG
met gln trp asn ser thr ala phe his gln ala leu gln asp pro arg val arg gly leu
61  /  21                                  91  /  31
TAT TTT CCT GCT GGT GGC TCC AGT TCA GGA ACA GTA AAC CCT GCT CCG AAT ATT GCC TCT
tyr phe pro ala gly gly ser ser ser gly thr val asn pro ala pro asn ile ala ser
121 /  41                                  151 /  51
CAC ATC TCG TCA ATC TCC GCG AGG ACT GGG GAC CCT GTG ACG AAC ATG GAG AAC ATC ACA
his ile ser ser ile ser ala arg thr gly asp pro val thr asn met glu asn ile thr
181 /  61                                  211 /  71
TCA GGA TTC CTA GGA CCC CTG CTC GTG TTA CAG GCG GGG TTT TTC TTG TTG ACA AGA ATC
ser gly phe leu gly pro leu leu val leu gln ala gly phe phe leu leu thr arg ile
241 /  81                                  271 /  91
CTC ACA ATA CCG CAG AGT CTA GAC TCG TGG TGG ACT TCT CTC AAT TTT CTA GGG GGA TCA
leu thr ile pro gln ser leu asp ser trp trp thr ser leu asn phe leu gly gly ser
301 / 101                                  331 / 111
CCC GTG TGT CTT GGC CAA AAT TCG CAG TCC CCA ACC TCC AAT CAC TCA CCA ACC TCC TGT
pro val cys leu gly gln asn ser gln ser pro thr ser asn his ser pro thr ser cys
361 / 121                                  391 / 131
CCT CCA ATT TGT CCT GGT TAT CGC TGG ATG TGT CTG CGG CGT TTT ATC ATA TTC CTC TTC
pro pro ile cys pro gly tyr arg trp met cys leu arg arg phe ile ile phe leu phe
421 / 141                                  451 / 151
ATC CTG CTG CTA TGC CTC ATC TTC TTA TTG GTT CTT CTG GAT TAT CAA GGT ATG TTG CCC
ile leu leu leu cys leu ile phe leu leu val leu leu asp tyr gln gly met leu pro
481 / 161                                  511 / 171
GTT TGT CCT CTA ATT CCA GGA TCA ACA ACA ACC AGT ACG GGA CCA TGC AAA ACC TGC ACG
val cys pro leu ile pro gly ser thr thr thr ser thr gly pro cys lys thr cys thr
541 / 181                                  571 / 191
ACT CCT GCT CAA GGC AAC TCT ATG TTT CCC TCA TGT TGC TGT ACA AAA CCT ACG GAT GGA
thr pro ala gln gly asn ser met phe pro ser cys cys cys thr lys pro thr asp gly
601 / 201                                  631 / 211
AAT TGC ACC TGT ATT CCC ATC CCA TCG TCT TGG GCT TTC GCA AAA TAC CTA TGG GAG TGG
asn cys thr cys ile pro ile pro ser ser trp ala phe ala lys tyr leu trp glu trp
661 / 221                                  691 / 231
GCC TCA GTC CGT TTC TCT TGG CTC AGT TTA CTA GTG CCA TTT GTT CAG TGG TTC GTA GGG
ala ser val arg phe ser trp leu ser leu leu val pro phe val gln trp phe val gly
721 / 241                                  751 / 251
CTT TCC CCC ACT GTT TGG CTT TCA GCT ATA TGG ATG ATG TGG TAT TGG GGG CCA AGT CTG
leu ser pro thr val trp leu ser ala ile trp met met trp tyr trp gly pro ser leu
781 / 261                                  811 / 271
TAC AGC ATC GTG AGT CCC TTT ATA CCG CTG TTA CCA ATT TTC TTT TGT CTC TGG GTA TAC
tyr ser ile val ser pro phe ile pro leu leu pro ile phe phe cys leu trp val tyr
841 / 281
ATT TAA
ile OCH
```

FIG.23

```
1   /   1                                31   /   11
ATG GAG AAC ATC ACA TCA GGA TTC CTA GGA CCC CTG CTC GTG TTA CAG GCG GGG TTT TTC
met glu asn ile thr ser gly phe leu gly pro leu leu val leu gln ala gly phe phe
61  /   21                               91   /   31
TTG TTG ACA AGA ATC CTC ACA ATA CCG CAG AGT CTA GAC TCG TGG TGG ACT TCT CTC AAT
leu leu thr arg ile leu thr ile pro gln ser leu asp ser trp trp thr ser leu asn
121 /   41                               151  /   51
TTT CTA GGG GGA TCA CCC GTG TGT CTT GGC CAA AAT TCG CAG TCC CCA ACC TCC AAT CAC
phe leu gly gly ser pro val cys leu gly gln asn ser gln ser pro thr ser asn his
181 /   61                               211  /   71
TCA CCA ACC TCC TGT CCT CCA ATT TGT CCT GGT TAT CGC TGG ATG TGT CTG CGG CGT TTT
ser pro thr ser cys pro pro ile cys pro gly tyr arg trp met cys leu arg arg phe
241 /   81                               271  /   91
ATC ATA TTC CTC TTC ATC CTG CTG CTA TGC CTC ATC TTC TTA TTG GTT CTT CTG GAT TAT
ile ile phe leu phe ile leu leu leu cys leu ile phe leu leu val leu leu asp tyr
301 /  101                               331  /  111
CAA GGT ATG TTG CCC GTT TGT CCT CTA ATT CCA GGA TCA ACA ACA ACC AGT ACG GGA CCA
gln gly met leu pro val cys pro leu ile pro gly ser thr thr thr ser thr gly pro
361 /  121                               391  /  131
TGC AAA ACC TGC ACG ACT CCT GCT CAA GGC AAC TCT ATG TTT CCC TCA TGT TGC TGT ACA
cys lys thr cys thr thr pro ala gln gly asn ser met phe pro ser cys cys cys thr
421 /  141                               451  /  151
AAA CCT ACG GAT GGA AAT TGC ACC TGT ATT CCC ATC CCA TCG TCT TGG GCT TTC GCA AAA
lys pro thr asp gly asn cys thr cys ile pro ile pro ser ser trp ala phe ala lys
481 /  161                               511  /  171
TAC CTA TGG GAG TGG GCC TCA GTC CGT TTC TCT TGG CTC AGT TTA CTA GTG CCA TTT GTT
tyr leu trp glu trp ala ser val arg phe ser trp leu ser leu leu val pro phe val
541 /  181                               571  /  191
CAG TGG TTC GTA GGG CTT TCC CCC ACT GTT TGG CTT TCA GCT ATA TGG ATG ATG TGG TAT
gln trp phe val gly leu ser pro thr val trp leu ser ala ile trp met met trp tyr
601 /  201                               631  /  211
TGG GGG CCA AGT CTG TAC AGC ATC GTG AGT CCC TTT ATA CCG CTG TTA CCA ATT TTC TTT
trp gly pro ser leu tyr ser ile val ser pro phe ile pro leu leu pro ile phe phe
661 /  221
TGT CTC TGG GTA TAC ATT TAA
cys leu trp val tyr ile OCH
```

FIG.24

FIG. 25

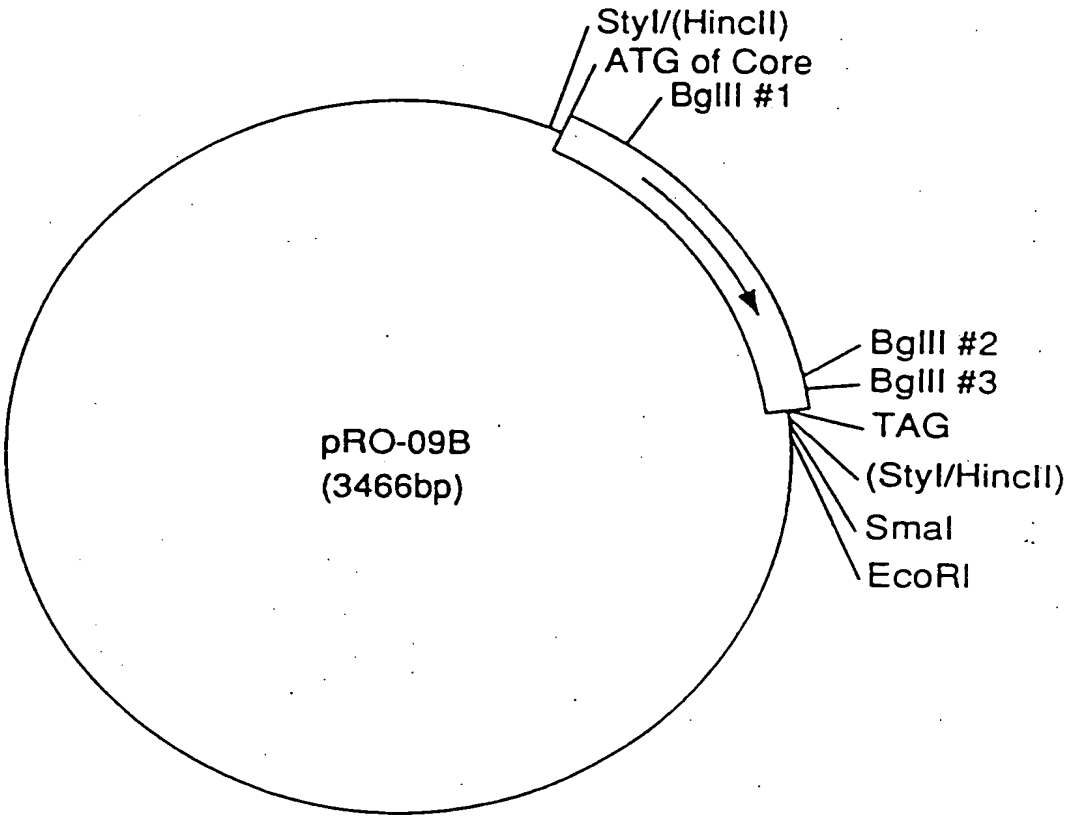

StyI/(HincII)
ATG of Core
BgIII #1
BgIII #2
BgIII #3
TAG
(StyI/HincII)
SmaI
EcoRI
pRO-09B
(3466bp)

FIG. 26

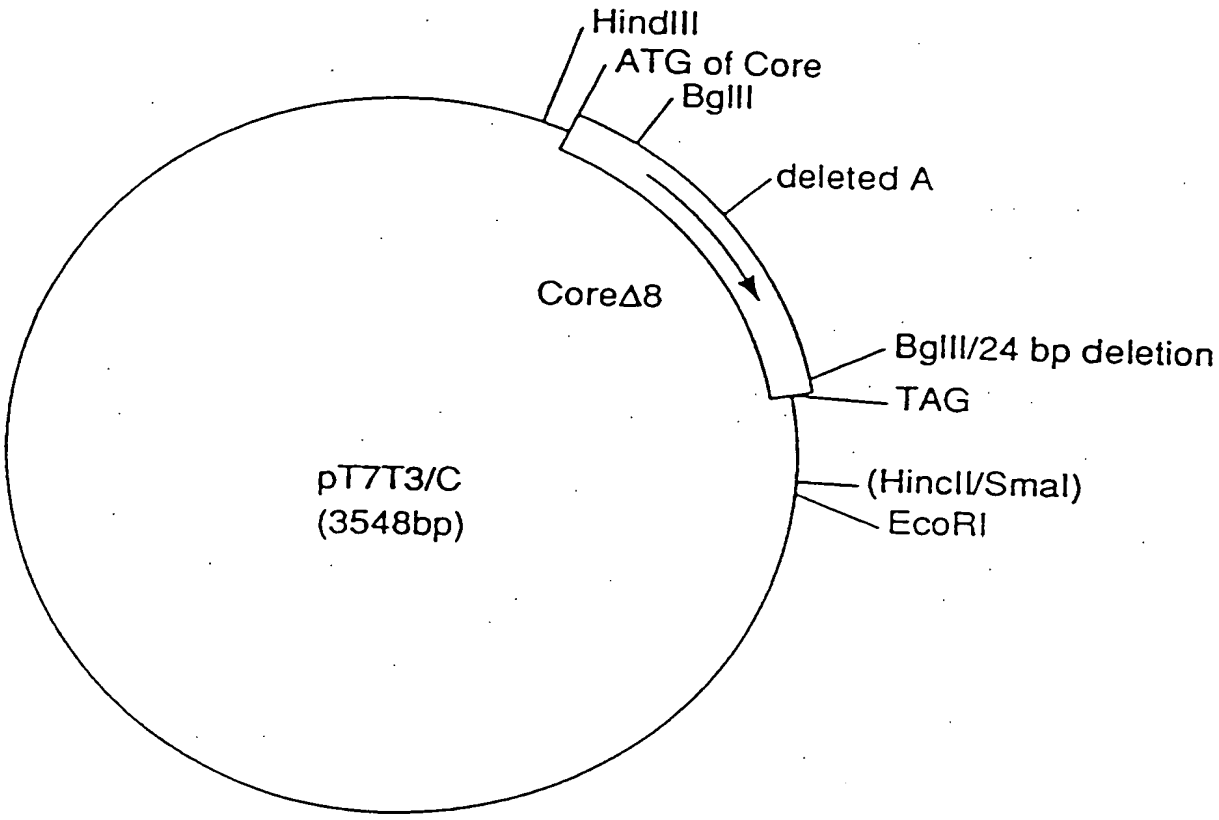

HindIII
ATG of Core
BglII
deleted A
CoreΔ8
BglII/24 bp deletion
TAG
pT7T3/C
(3548bp)
(HincII/SmaI)
EcoRI

FIG. 27

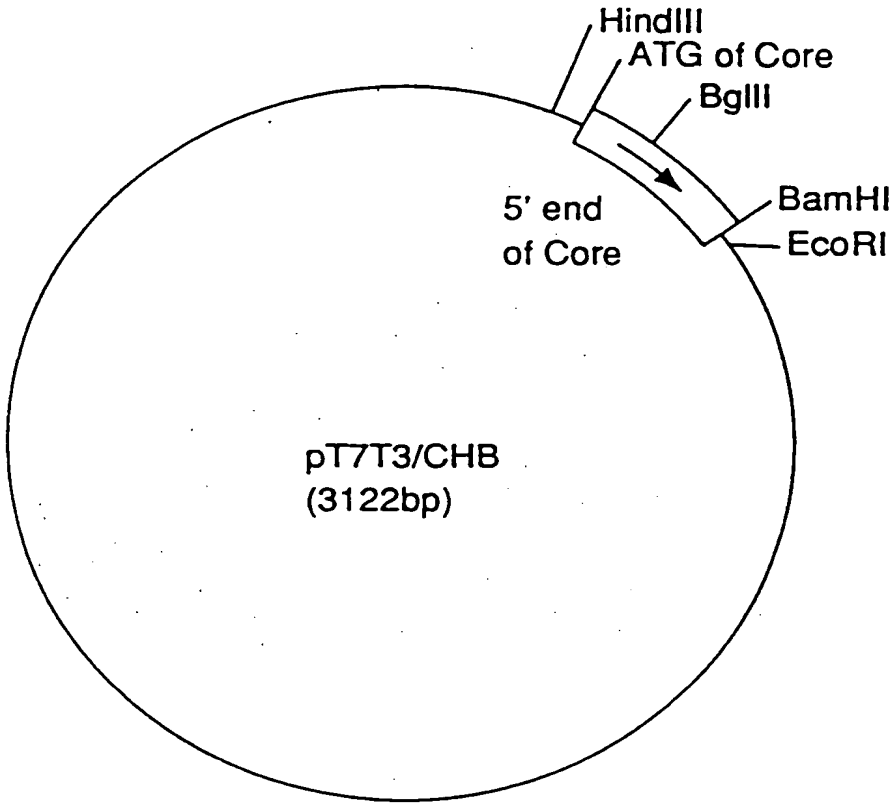

HindIII
ATG of Core
BglII
BamHI
EcoRI
5' end
of Core
pT7T3/CHB
(3122bp)

FIG. 28

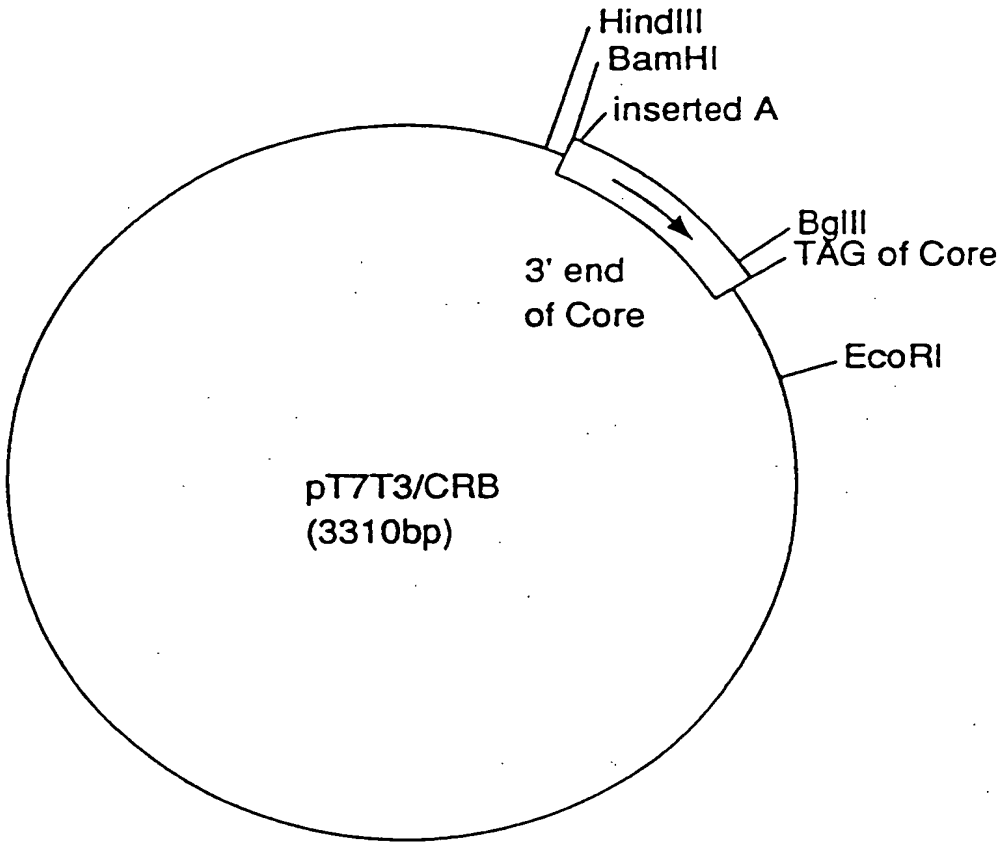
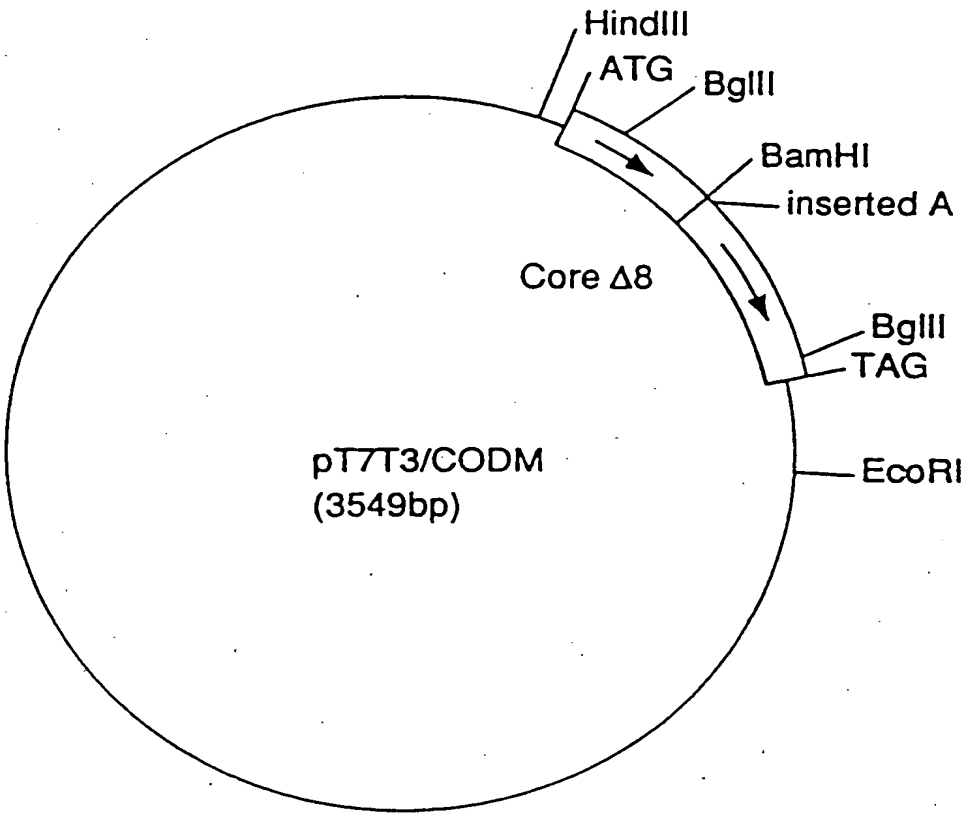

HindIII
BamHI
inserted A
BglII
TAG of Core
3' end
of Core
EcoRI
pT7T3/CRB
(3310bp)

FIG. 29


HindIII
ATG
BglII
BamHI
inserted A
Core Δ8
BglII
TAG
pT7T3/CODM
(3549bp)
EcoRI

```
1   /   1                                31   /   11
ATG GAC ATT GAC CCT TAT AAA GAA TTT GGA GCT ACT GTG GAG TTA CTC TCG TTT TTG CCT
met asp ile asp pro tyr lys glu phe gly ala thr val glu leu leu ser phe leu pro
61   /   21                              91   /   31
TCT GAC TTC TTT CCT TCC GTC AGA GAT CTC CTA GAC ACC GCC TCA GCT CTG TAT CGG GAA
ser asp phe phe pro ser val arg asp leu leu asp thr ala ser ala leu tyr arg glu
121   /   41                             151   /   51
GCC TTA GAG TCT CCT GAG CAT TGC TCA CCT CAC CAC ACC GCA CTC AGG CAA GCC ATT CTC
ala leu glu ser pro glu his cys ser pro his his thr ala leu arg gln ala ile leu
181   /   61                             211   /   71
TGC TGG GGG GAA TTG ATG ACT CTA GCT ACC TGG GTG GGT AAT AAT TTG GAG GAT CCA GCA
cys trp gly glu leu met thr leu ala thr trp val gly asn asn leu glu asp pro ala
241   /   81                             271   /   91
TCA AGG GAT CTA GTA GTC AAT TAT GTT AAT ACT AAC ATG GGT TTA AAA ATT AGG CAA CTA
ser arg asp leu val val asn tyr val asn thr asn met gly leu lys ile arg gln leu
301   /   101                            331   /   111
TTG TGG TTT CAT ATA TCT TGC CTT ACT TTT GGA AGA GAG ACT GTA CTT GAA TAT TTG GTA
leu trp phe his ile ser cys leu thr phe gly arg glu thr val leu glu tyr leu val
361   /   121                            391   /   131
TCT TTC GGA GTG TGG ATT CGC ACT CCT CCA GCC TAT AGA CCA CCA AAT GCC CCT ATC TTA
ser phe gly val trp ile arg thr pro pro ala tyr arg pro pro asn ala pro ile leu
421   /   141                            451   /   151
TCA ACA CTT CCG GAA ACT ACT GTT GTT AGA CGA CGG GAC CGA GGC AGG TCC CCT AGA AGA
ser thr leu pro glu thr thr val val arg arg arg asp arg gly arg ser pro arg arg
481   /   161                            511   /   171
AGA ACT CCC TCG CCT CGC AGA CGC AGA TCT CAA TCT CGG GAA TCT CAA TGT TAG
arg thr pro ser pro arg arg arg arg ser gln ser arg glu ser gln cys AMB
```

FIG.30

FIG. 31

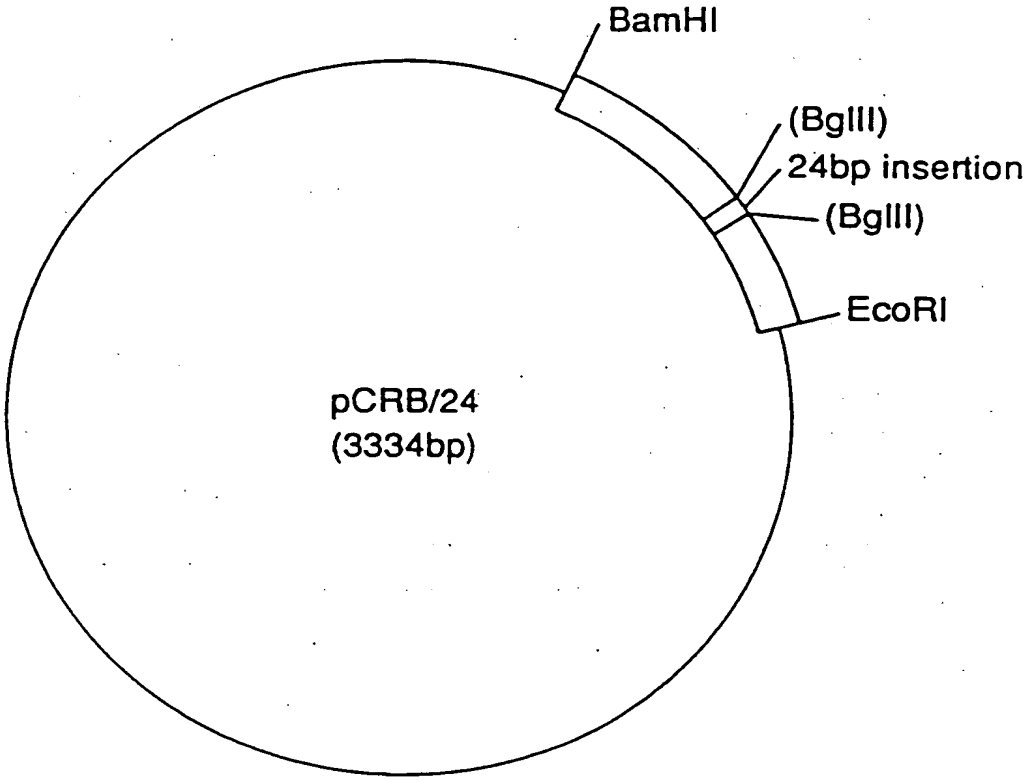
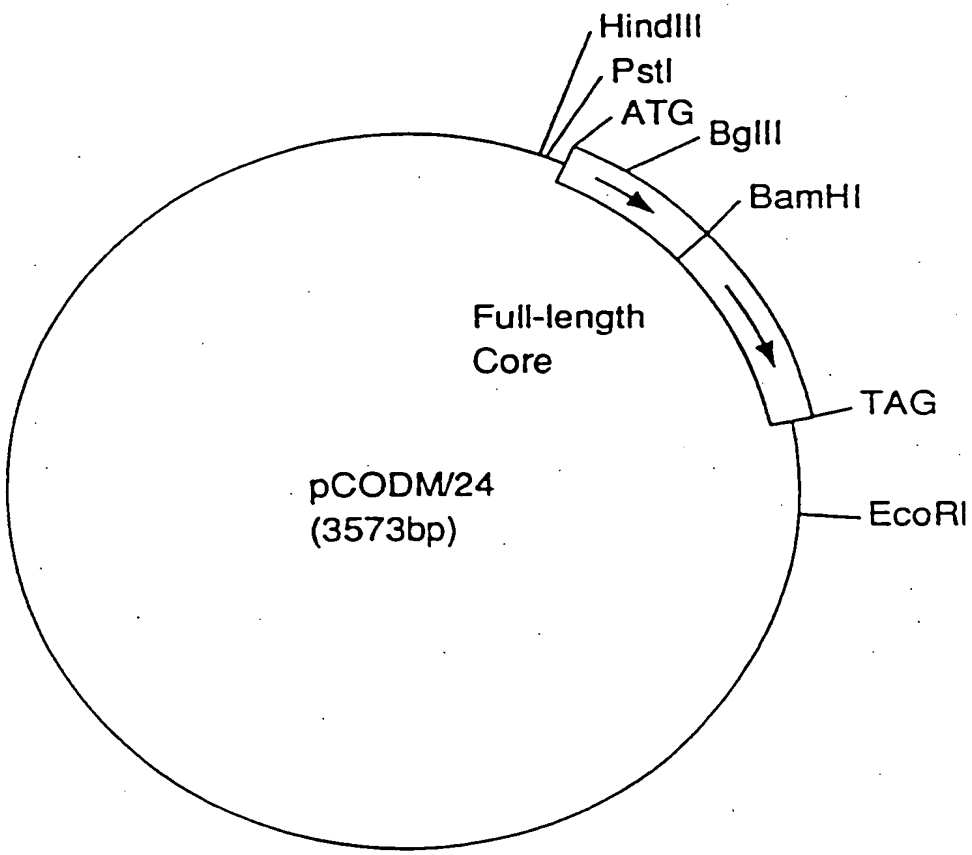


FIG. 32


HindIII
PstI
ATG
BglII
BamHI
Full-length
Core
TAG
pCODM/24
(3573bp)
EcoRI

```
1   /   1                                31   /   11
ATG GAC ATT GAC CCT TAT AAA GAA TTT GGA GCT ACT GTG GAG TTA CTC TCG TTT TTG CCT
met asp ile asp pro tyr lys glu phe gly ala thr val glu leu leu ser phe leu pro
61   /   21                              91   /   31
TCT GAC TTC TTT CCT TCC GTC AGA GAT CTC CTA GAC ACC GCC TCA GCT CTG TAT CGG GAA
ser asp phe phe pro ser val arg asp leu leu asp thr ala ser ala leu tyr arg glu
121  /   41                              151  /   51
GCC TTA GAG TCT CCT GAG CAT TGC TCA CCT CAC CAC ACC GCA CTC AGG CAA GCC ATT CTC
ala leu glu ser pro glu his cys ser pro his his thr ala leu arg gln ala ile leu
181  /   61                              211  /   71
TGC TGG GGG GAA TTG ATG ACT CTA GCT ACC TGG GTG GGT AAT AAT TTG GAG GAT CCA GCA
cys trp gly glu leu met thr leu ala thr trp val gly asn asn leu glu asp pro ala
241  /   81                              271  /   91
TCA AGG GAT CTA GTA GTC AAT TAT GTT AAT ACT AAC ATG GGT TTA AAA ATT AGG CAA CTA
ser arg asp leu val val asn tyr val asn thr asn met gly leu lys ile arg gln leu
301  /  101                             331  /  111
TTG TGG TTT CAT ATA TCT TGC CTT ACT TTT GGA AGA GAG ACT GTA CTT GAA TAT TTG GTA
leu trp phe his ile ser cys leu thr phe gly arg glu thr val leu glu tyr leu val
361  /  121                             391  /  131
TCT TTC GGA GTG TGG ATT CGC ACT CCT CCA GCC TAT AGA CCA CCA AAT GCC CCT ATC TTA
ser phe gly val trp ile arg thr pro pro ala tyr arg pro pro asn ala pro ile leu
421  /  141                             451  /  151
TCA ACA CTT CCG GAA ACT ACT GTT GTT AGA CGA CGG GAC CGA GGC AGG TCC CCT AGA AGA
ser thr leu pro glu thr thr val val arg arg arg asp arg gly arg ser pro arg arg
481  /  161                             511  /  171
AGA ACT CCC TCG CCT CGC AGA CGC AGA tcc caa tcg ccg cgt cgc aga cga TCT CAA TCT
aga thr pro ser pro arg arg arg arg ser gln ser pro arg arg arg arg ser gln ser
541  /  181
CGG GAA TCT CAA TGT TAG
arg glu ser gln cys AMB
```

FIG.33

FIG. 34

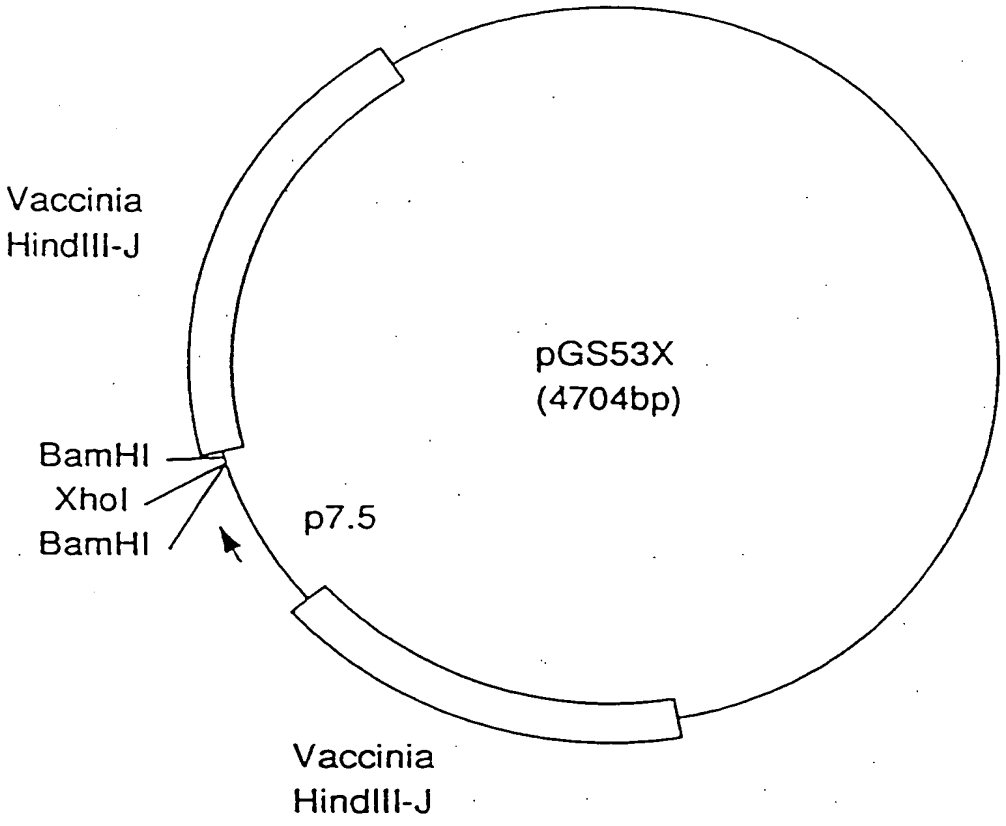

Vaccinia
HindIII-J
pGS53X
(4704bp)
BamHI
XhoI
BamHI
p7.5
Vaccinia
HindIII-J

FIG. 35

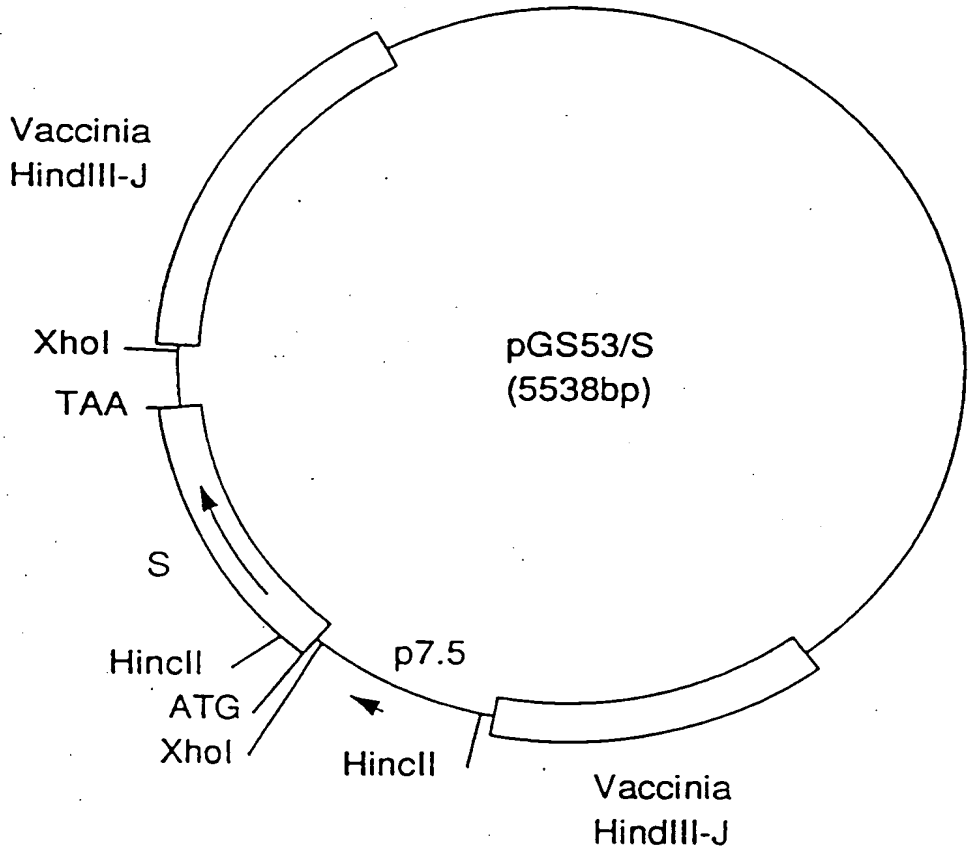

Vaccinia
HindIII-J
XhoI
TAA
pGS53/S
(5538bp)
S
HincII
ATG
XhoI
p7.5
HincII
Vaccinia
HindIII-J

FIG. 36

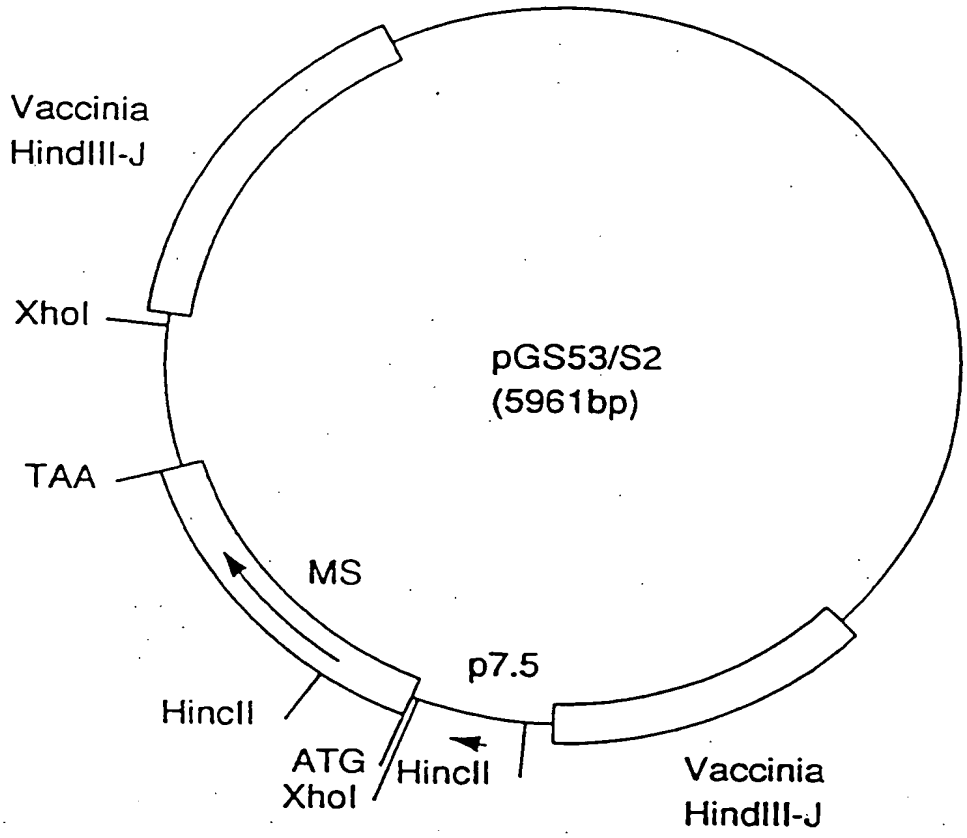

Vaccinia
HindIII-J
XhoI
TAA
MS
HincII
ATG
XhoI
p7.5
HincII
pGS53/S2
(5961bp)
Vaccinia
HindIII-J

FIG. 37

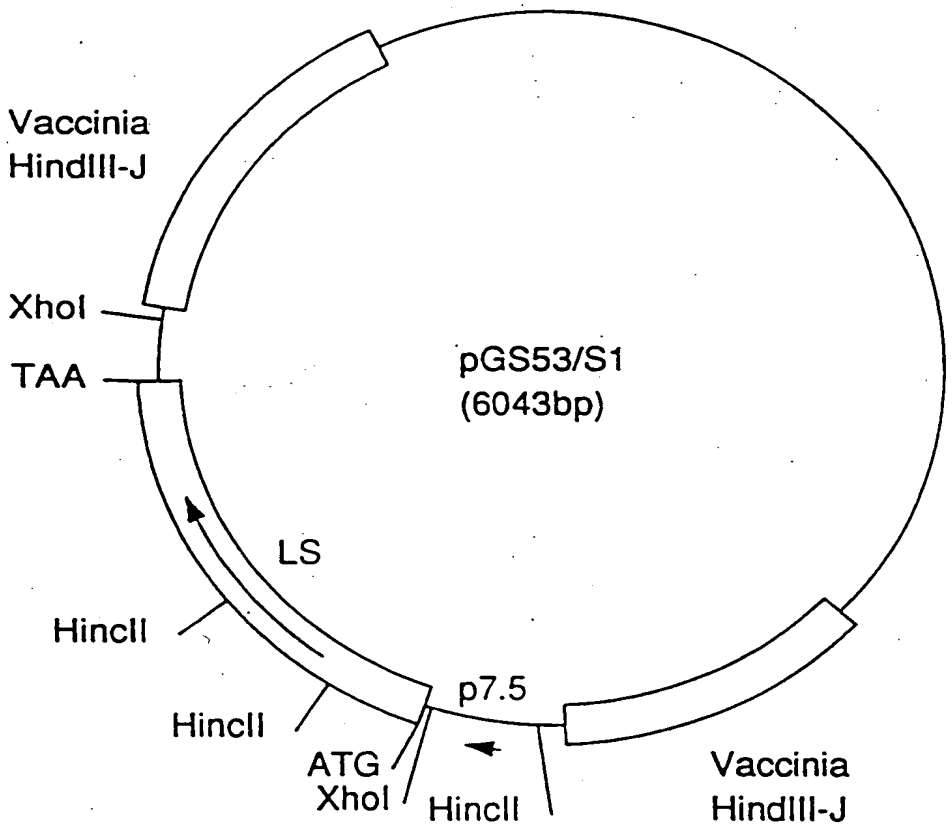

Vaccinia
HindIII-J
XhoI
TAA
pGS53/S1
(6043bp)
LS
HincII
HincII
p7.5
ATG
XhoI
HincII
Vaccinia
HindIII-J

FIG. 38

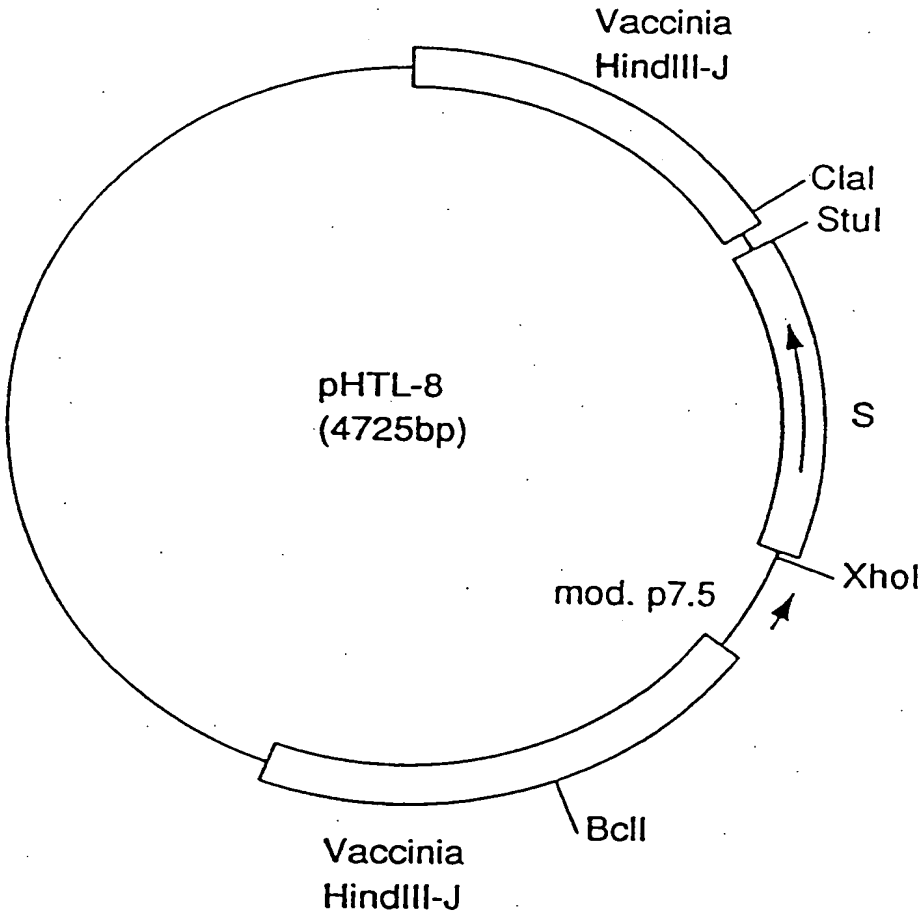

Vaccinia
HindIII-J
ClaI
StuI
pHTL-8
(4725bp)
S
XhoI
mod. p7.5
BclI
Vaccinia
HindIII-J

FIG. 39

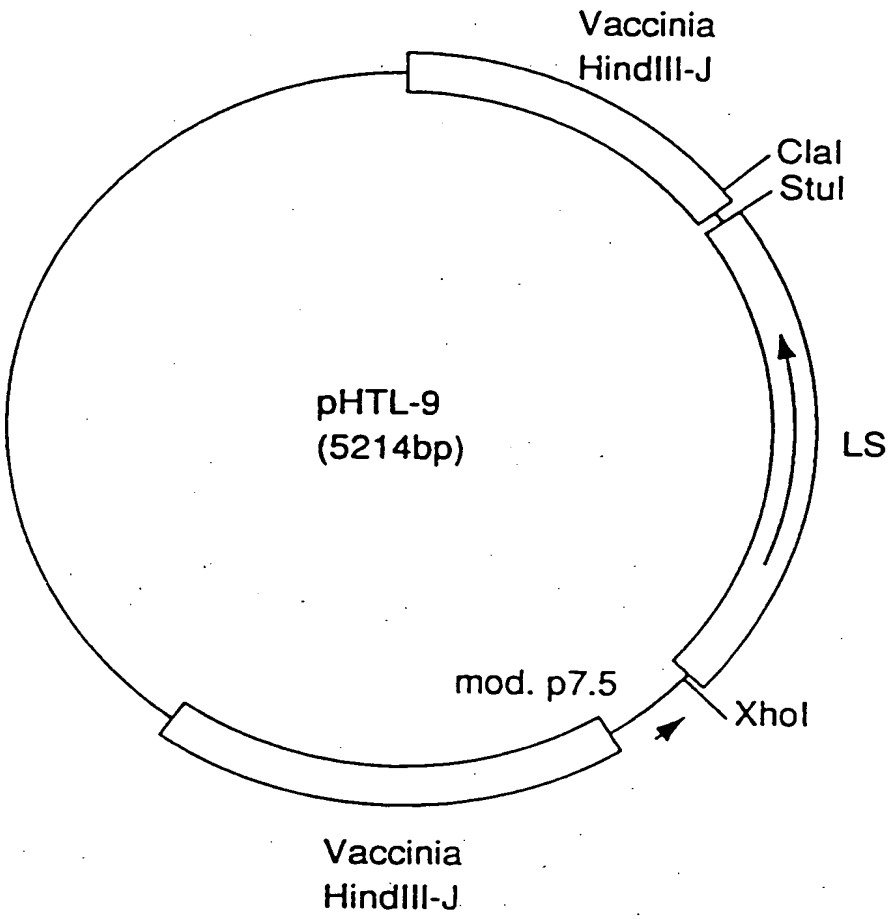

Vaccinia
HindIII-J
ClaI
StuI
pHTL-9
(5214bp)
LS
mod. p7.5
XhoI
Vaccinia
HindIII-J

FIG. 40

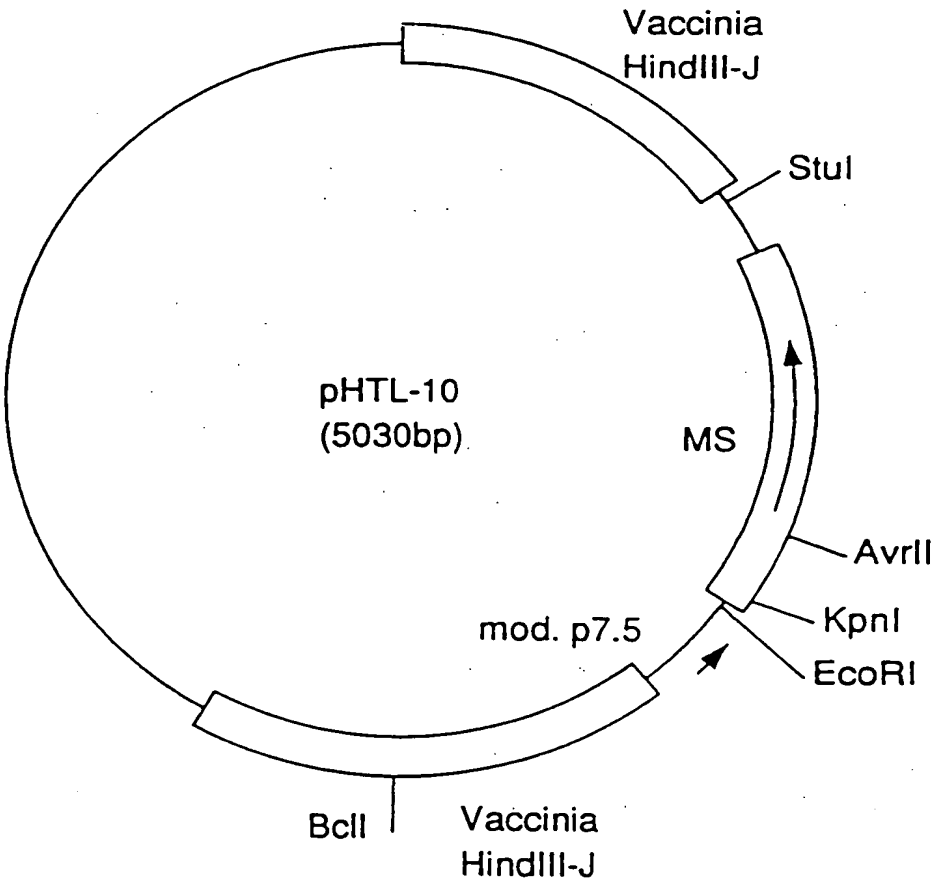

Vaccinia
HindIII-J
StuI
pHTL-10
(5030bp)
MS
AvrII
KpnI
mod. p7.5
EcoRI
BclI
Vaccinia
HindIII-J

FIG. 41

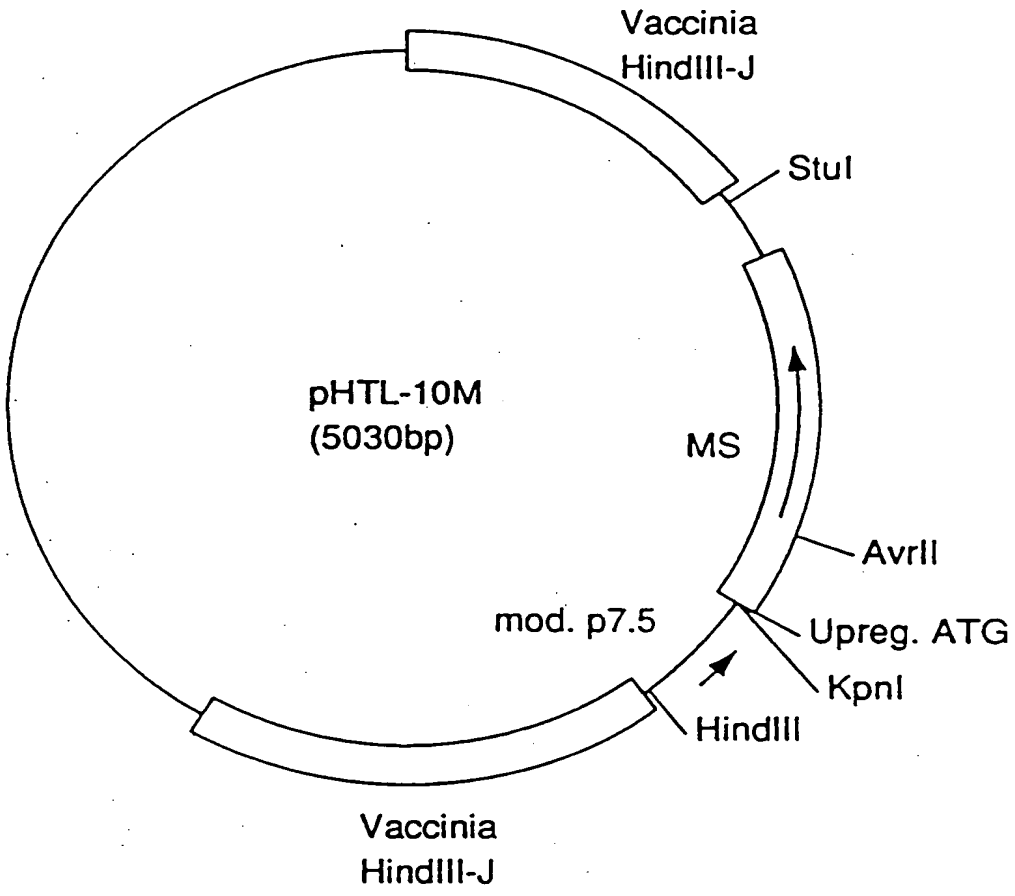

Vaccinia
HindIII-J
StuI
pHTL-10M
(5030bp)
MS
AvrII
mod. p7.5
Upreg. ATG
KpnI
HindIII
Vaccinia
HindIII-J

FIG. 42

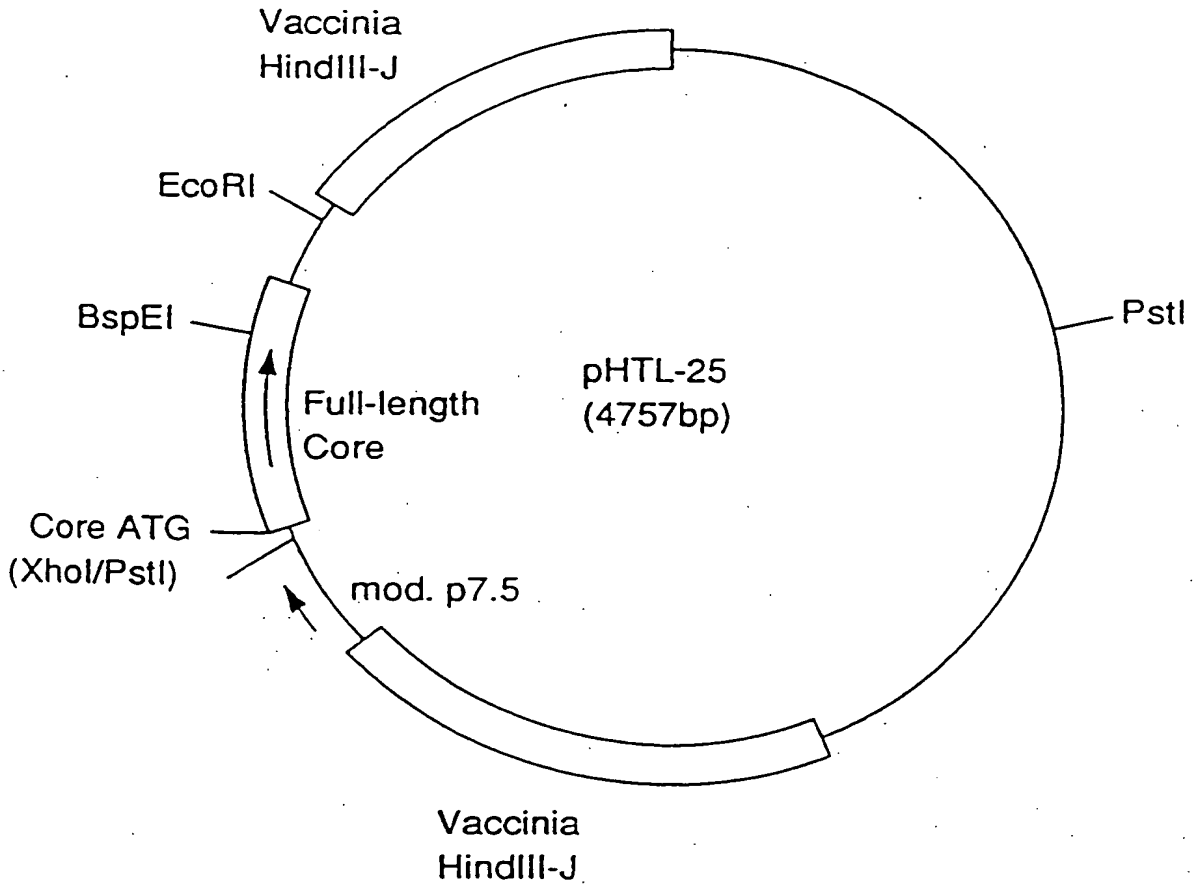
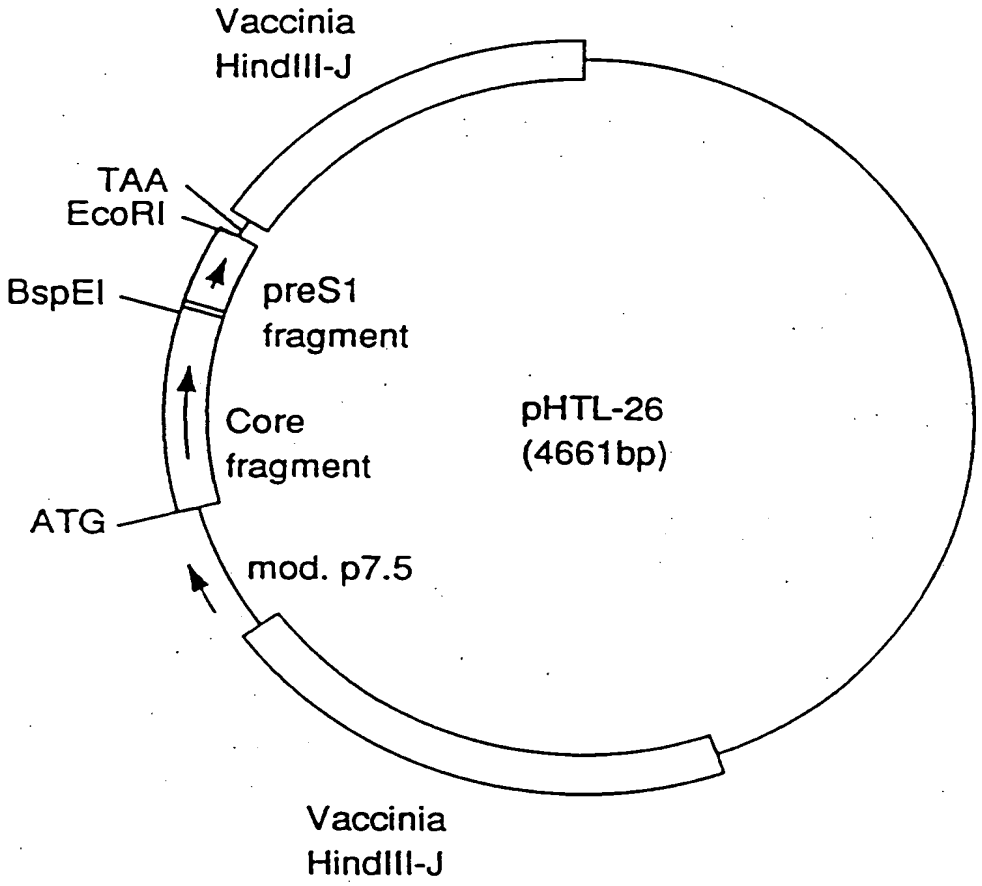

Vaccinia
HindIII-J
EcoRI
BspEI
Full-length
Core
pHTL-25
(4757bp)
PstI
Core ATG
(XhoI/PstI)
mod. p7.5
Vaccinia
HindIII-J

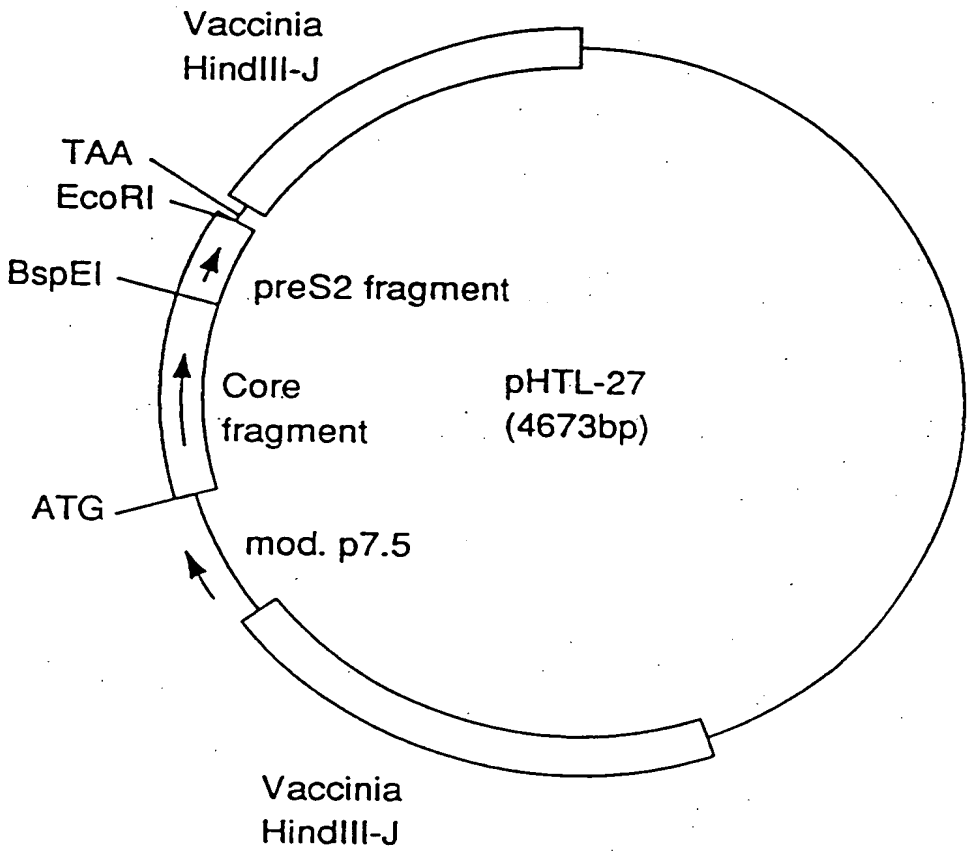
FIG. 43


Vaccinia
HindIII-J
TAA
EcoRI
BspEI
preS1
fragment
Core
fragment
ATG
mod. p7.5
pHTL-26
(4661bp)
Vaccinia
HindIII-J

```
1   /   1                                   31   /   11
ATG GAC ATT GAC CCT TAT AAA GAA TTT GGA GCT ACT GTG GAG TTA CTC TCG TTT TTG CCT
met asp ile asp pro tyr lys glu phe gly ala thr val glu leu leu ser phe leu pro
61   /   21                                 91   /   31
TCT GAC TTC TTT CCT TCC GTC AGA GAT CTC CTA GAC ACC GCC TCA GCT CTG TAT CGG GAA
ser asp phe phe pro ser val arg asp leu leu asp thr ala ser ala leu tyr arg glu
121  /   41                                 151  /   51
GCC TTA GAG TCT CCT GAG GAT TGC TCA CCT CAC CAC ACC GCA CTC AGG CAA GCC ATT CTC
ala leu glu ser pro glu his cys ser pro his his thr ala leu arg gln ala ile leu
181  /   61                                 211  /   71
TGC TGG GGG GAA TTG ATG ACT CTA GCT ACC TGG GTG GGT AAT AAT TTG GAG GAT CCA GCA
cys trp gly glu leu met thr leu ala thr trp val gly asn asn leu glu asp pro ala
241  /   81                                 271  /   91
TCA AGG GAT CTA GTA GTC AAT TAT GTT AAT ACT AAC ATG GGT TTA AAA ATT AGG CAA CTA
ser arg asp leu val val asn tyr val asn thr asn met gly leu lys ile arg gln leu
301  /  101                                 331  /  111
TTG TGG TTT CAT ATA TCT TGC CTT ACT TTT GGA AGA GAG ACT GTA CTT GAA TAT TTG GTA
leu trp phe his ile ser cys leu thr phe gly arg glu thr val leu glu tyr leu val
361  /  121                                 391  /  131
TCT TTC GGA GTG TGG ATT CGC ACT CCT CCA GCC TAT AGA CCA CCA AAT GCC CCT ATC TTA
ser phe gly val trp ile arg thr pro pro ala tyr arg pro pro asn ala pro ile leu
421  /  141                                 451  /  151
TCA ACA CTT CCg gaa tCa gct tgc ATG GGG ACG AAT CTT TCT GTT CCC AAt CCT CTG GGA
ser thr leu pro glu ser ala cys met gly thr asn leu ser val pro asn pro leu gly
481  /  161                                 511  /  171
TTC tTT CCC GAT CAT CAG TTG GAC CCT GCA TTC GGA GCC AAC TCA Aac AAT CCA GAT TGG
phe phe pro asp his gln leu asp pro ala phe gly ala asn ser asn asn pro asp trp
541  /  181                                 571  /  191
GAC TTC AAC CCC ATC AAG GAC CAC TGG CCA GCA GCC AAC CAG GTA GGA GTG GGA GCA TTC
asp phe asn pro ile lys asp his trp pro ala ala asn gln val gly val gly ala phe
601  /  201
GGG CCA GAA TTC AGG CCT ACT AGT TAA
gly pro glu phe arg pro thr ser OCH
```

FIG.44

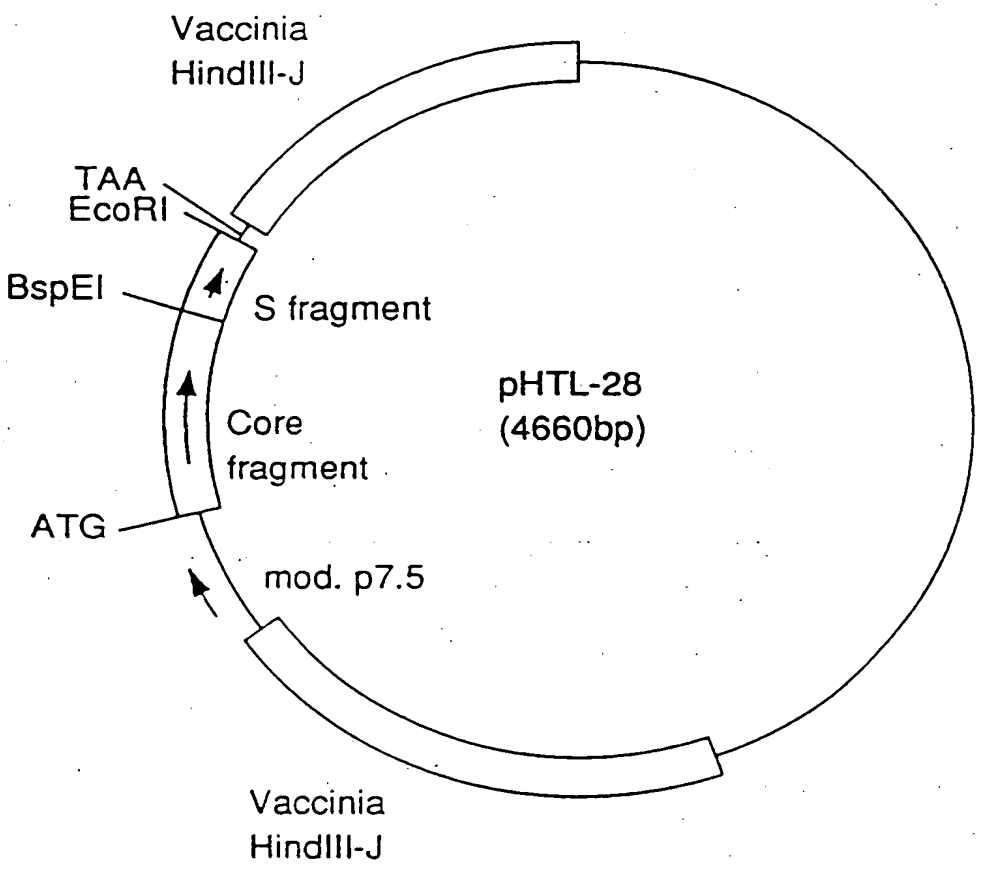
FIG. 45


Vaccinia
HindIII-J
TAA
EcoRI
BspEI
preS2 fragment
Core
fragment
pHTL-27
(4673bp)
ATG
mod. p7.5
Vaccinia
HindIII-J

```
1   /   1                                31   /   11
ATG GAC ATT GAC CCT TAT AAA GAA TTT GGA GCT ACT GTG GAG TTA CTC TCG TTT TTG CCT
met asp ile asp pro tyr lys glu phe gly ala thr val glu leu leu ser phe leu pro
61  /   21                               91   /   31
TCT GAC TTC TTT CCT TCC GTC AGA GAT CTC CTA GAC ACC GCC TCA GCT CTG TAT CGG GAA
ser asp phe phe pro ser val arg asp leu leu asp thr ala ser ala leu tyr arg glu
121 /   41                               151  /   51
GCC TTA GAG TCT CCT GAG CAT TGC TCA CCT CAC CAC ACC GCA CTC AGG CAA GCC ATT CTC
ala leu glu ser pro glu his cys ser pro his his thr ala leu arg gln ala ile leu
181 /   61                               211  /   71
TGC TGG GGG GAA TTG ATG ACT CTA GCT ACC TGG GTG GGT AAT AAT TTG GAG GAT CCA GCA
cys trp gly glu leu met thr leu ala thr trp val gly asn asn leu glu asp pro ala
241 /   81                               271  /   91
TCA AGG GAT CTA GTA GTC AAT TAT GTT AAT ACT AAC ATG GGT TTA AAA ATT AGG CAA CTA
ser arg asp leu val val asn tyr val asn thr asn met gly leu lys ile arg gln leu
301 /   101                              331  /   111
TTG TGG TTT CAT ATA TCT TGC CTT ACT TTT GGA AGA GAG ACT GTA CTT GAA TAT TTG GTA
leu trp phe his ile ser cys leu thr phe gly arg glu thr val leu glu tyr leu val
361 /   121                              391  /   131
TCT TTC GGA GTG TGG ATT CGC ACT CCT CCA GCC TAT AGA CCA CCA AAT GCC CCT ATC TTA
ser phe gly val trp ile arg thr pro pro ala tyr arg pro pro asn ala pro ile leu
421 /   141                              451  /   151
TCA ACA CTT CCG GAC ATG CAG TGG AAc TCC ACT GCC TTc CAC CAA GCT CTG CAG GAT CCC
ser thr leu pro asp met gln trp asn ser thr ala phe his gln ala leu gln asp pro
481 /   161                              511  /   171
AGA GTC AGG GGT CTG TAT tTT CCT GCT GGT GGC TCC AGT TCA GGA ACA GTA AAC CCT GCT
arg val arg gly leu tyr phe pro ala gly gly ser ser ser gly thr val asn pro ala
541 /   181                              571  /   191
CCG AAT ATT GCC TCT CAC ATC TCG TCA ATC TCC GCG AGG ACT GGG GAC CCT GTG ACG AaC
pro asn ile ala ser his ile ser ser ile ser ala arg thr gly asp pro val thr asn
601 /   201                              631  /   211
ATG GAG AAC ATC ACA TCA GAA TTC AGG CCT ACT AGT TAA
met glu asn ile thr ser glu phe arg pro thr ser OCH
```

FIG.46

FIG. 47


Vaccinia
HindIII-J
TAA
EcoRI
BspEI
S fragment
pHTL-28
(4660bp)
Core
fragment
ATG
mod. p7.5
Vaccinia
HindIII-J

```
1   /   1                                31   /   11
ATG GAC ATT GAC CCT TAT AAA GAA TTT GGA GCT ACT GTG GAG TTA CTC TCG TTT TTG CCT
met asp ile asp pro tyr lys glu phe gly ala thr val glu leu leu ser phe leu pro
61   /   21                              91   /   31
TCT GAC TTC TTT CCT TCC GTC AGA GAT CTC CTA GAC ACC GCC TCA GCT CTG TAT CGG GAA
ser asp phe phe pro ser val arg asp leu leu asp thr ala ser ala leu tyr arg glu
121   /   41                             151   /   51
GCC TTA GAG TCT CCT GAG CAT TGC TCA CCT CAC CAC ACC GCA CTC AGG CAA GCC ATT CTC
ala leu glu ser pro glu his cys ser pro his his thr ala leu arg gln ala ile leu
181   /   61                             211   /   71
TGC TGG GGG GAA TTG ATG ACT CTA GCT ACC TGG GTG GGT AAT AAT TTG GAG GAT CCA GCA
cys trp gly glu leu met thr leu ala thr trp val gly asn asn leu glu asp pro ala
241   /   81                             271   /   91
TCA AGG GAT CTA GTA GTC AAT TAT GTT AAT ACT AAC ATG GGT TTA AAA ATT AGG CAA CTA
ser arg asp leu val val asn tyr val asn thr asn met gly leu lys ile arg gln leu
301   /   101                            331   /   111
TTG TGG TTT CAT ATA TCT TGC CTT ACT TTT GGA AGA GAG ACT GTA CTT GAA TAT TTG GTA
leu trp phe his ile ser cys leu thr phe gly arg glu thr val leu glu tyr leu val
361   /   121                            391   /   131
TCT TTC GGA GTG TGG ATT CGC ACT CCT CCA GCC TAT AGA CCA CCA AAT GCC CCT ATC TTA
ser phe gly val trp ile arg thr pro pro ala tyr arg pro pro asn ala pro ile leu
421   /   141                            451   /   151
TCA ACA CTT CCG GAT TGT CCT CTA ATT CCA GGA TCA ACA ACA ACC AGT ACG GGA CCA TGC
ser thr leu pro asp cys pro leu ile pro gly ser thr thr thr ser thr gly pro cys
481   /   161                            511   /   171
AAA ACC TGC ACG ACT CCT GCT CAA GGC AAC TCT ATG TTT CCC TCA TGT TGC TGT ACA AAA
lys thr cys thr thr pro ala gln gly asn ser met phe pro ser cys cys cys thr lys
541   /   181                            571   /   191
CCT ACG GAT GGA AAT TGC ACC TGT ATT CCC ATC CCA TCG TCt TGG GCT TTC GCA AAA TAC
pro thr asp gly asn cys thr cys ile pro ile pro ser ser trp ala phe ala lys tyr
601   /   201
CTA TGG AAT TCA GGC CTA CTA GTT AAG TAA
leu trp asn ser gly leu leu val lys
```

FIG.48

FIG. 49

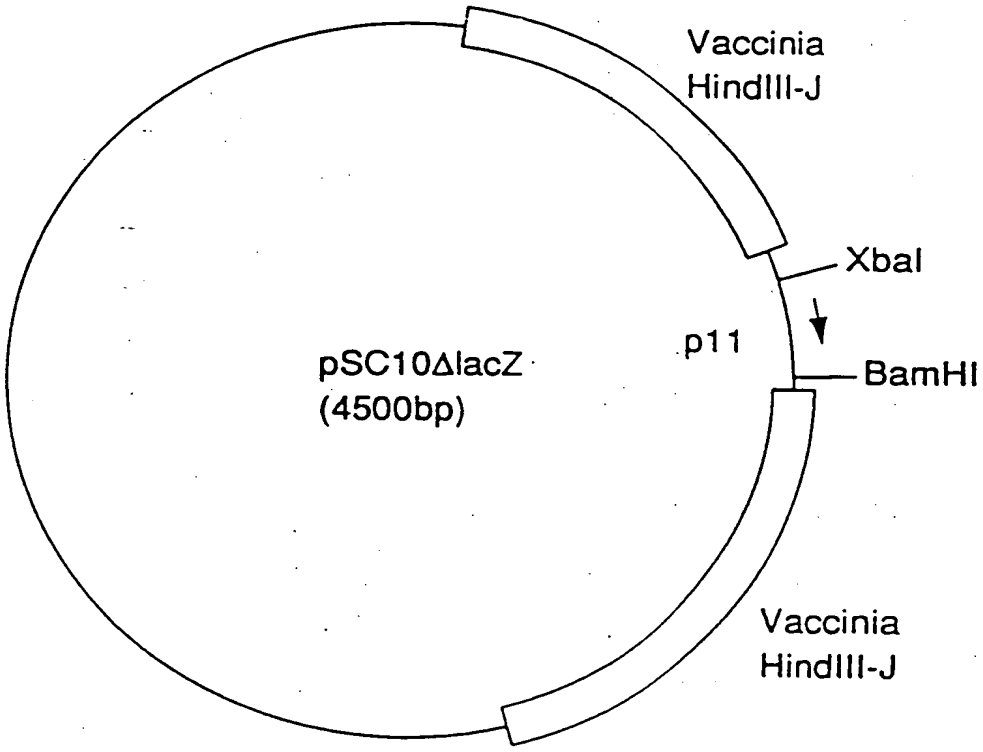

Vaccinia
HindIII-J
XbaI
p11
BamHI
pSC10ΔlacZ
(4500bp)
Vaccinia
HindIII-J

FIG. 50

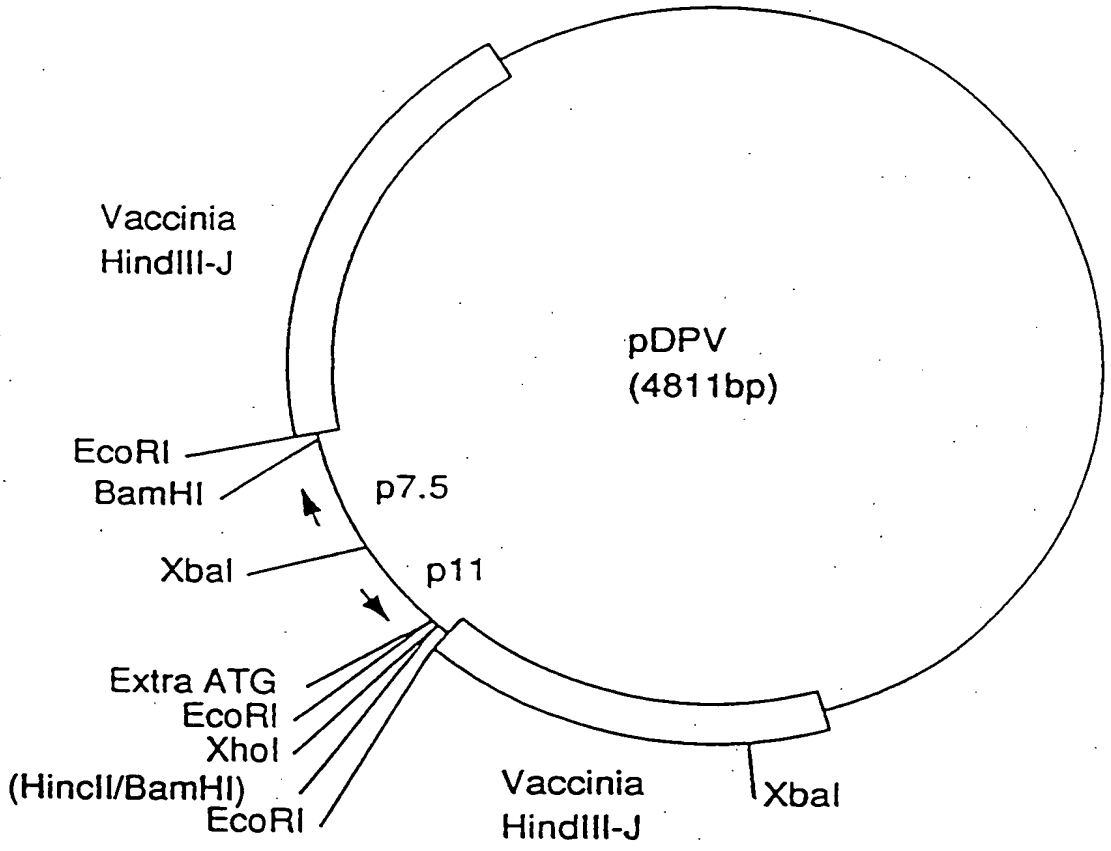

Vaccinia
HindIII-J
pDPV
(4811bp)
EcoRI
BamHI
p7.5
XbaI
p11
Extra ATG
EcoRI
XhoI
(HincII/BamHI)
EcoRI
Vaccinia
HindIII-J
XbaI

FIG. 51

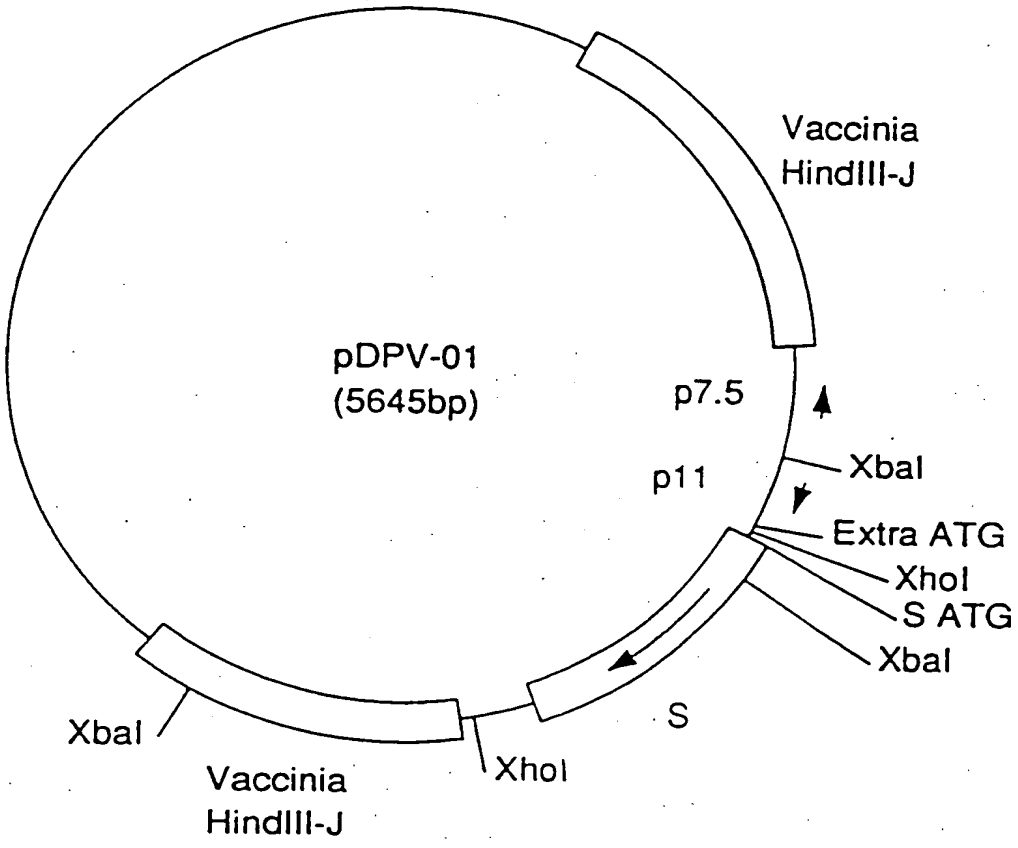

Vaccinia
HindIII-J
pDPV-01
(5645bp)
p7.5
p11
XbaI
Extra ATG
XhoI
S ATG
XbaI
S
XhoI
XbaI
Vaccinia
HindIII-J

FIG. 52

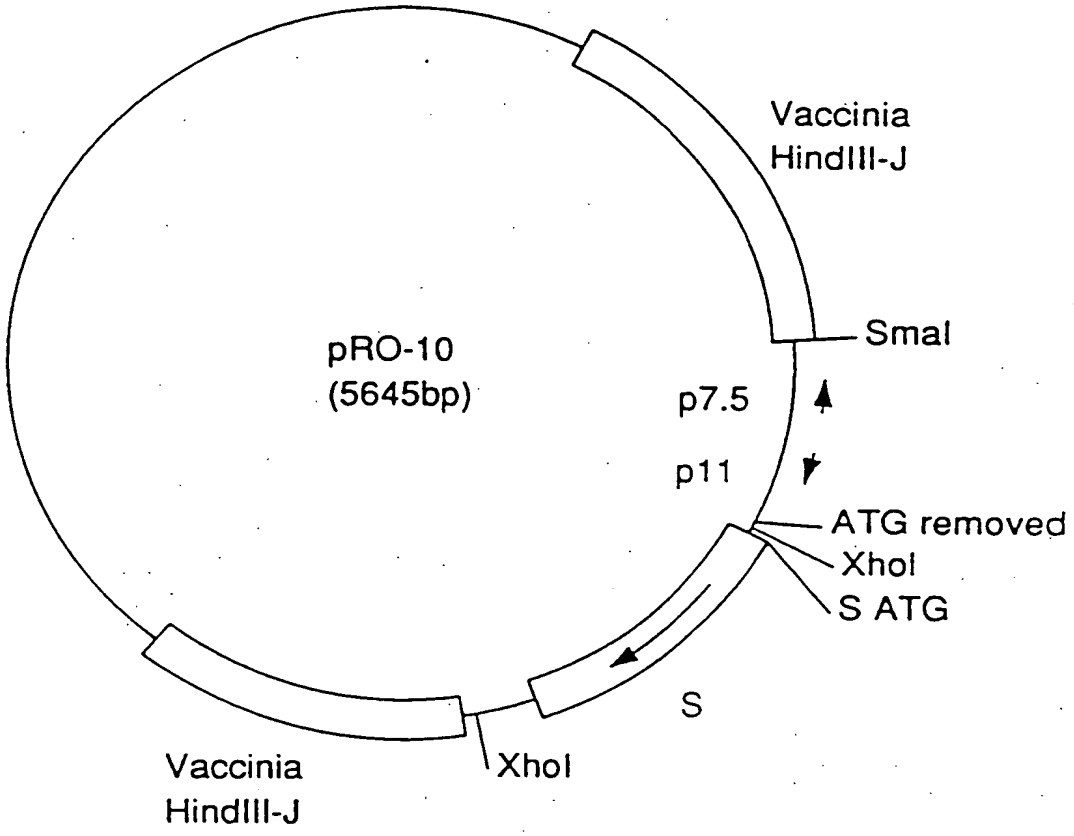

Vaccinia
HindIII-J
SmaI
pRO-10
(5645bp)
p7.5
p11
ATG removed
XhoI
S ATG
S
Vaccinia
HindIII-J
XhoI

FIG. 53

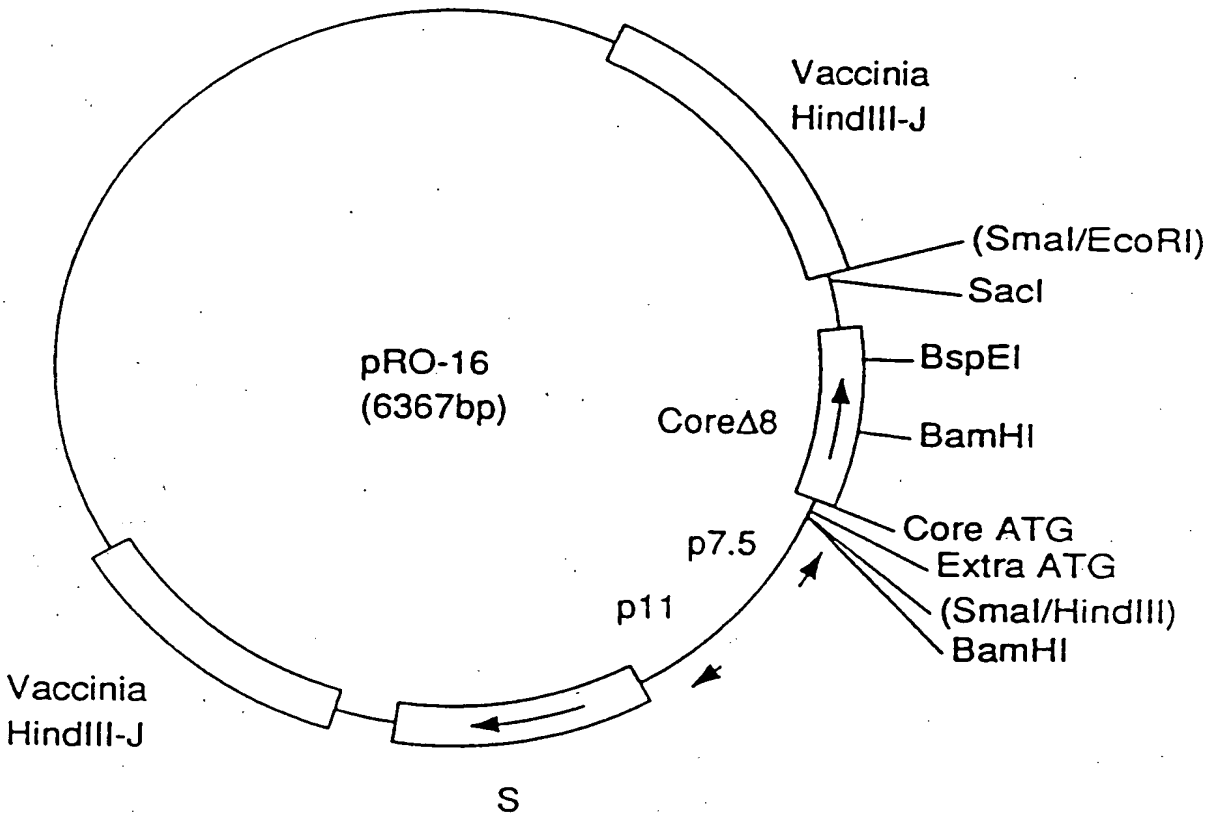

Vaccinia
HindIII-J
(SmaI/EcoRI)
SacI
BspEI
BamHI
Core ATG
Extra ATG
(SmaI/HindIII)
BamHI
pRO-16
(6367bp)
CoreΔ8
p7.5
p11
Vaccinia
HindIII-J
S

FIG. 54

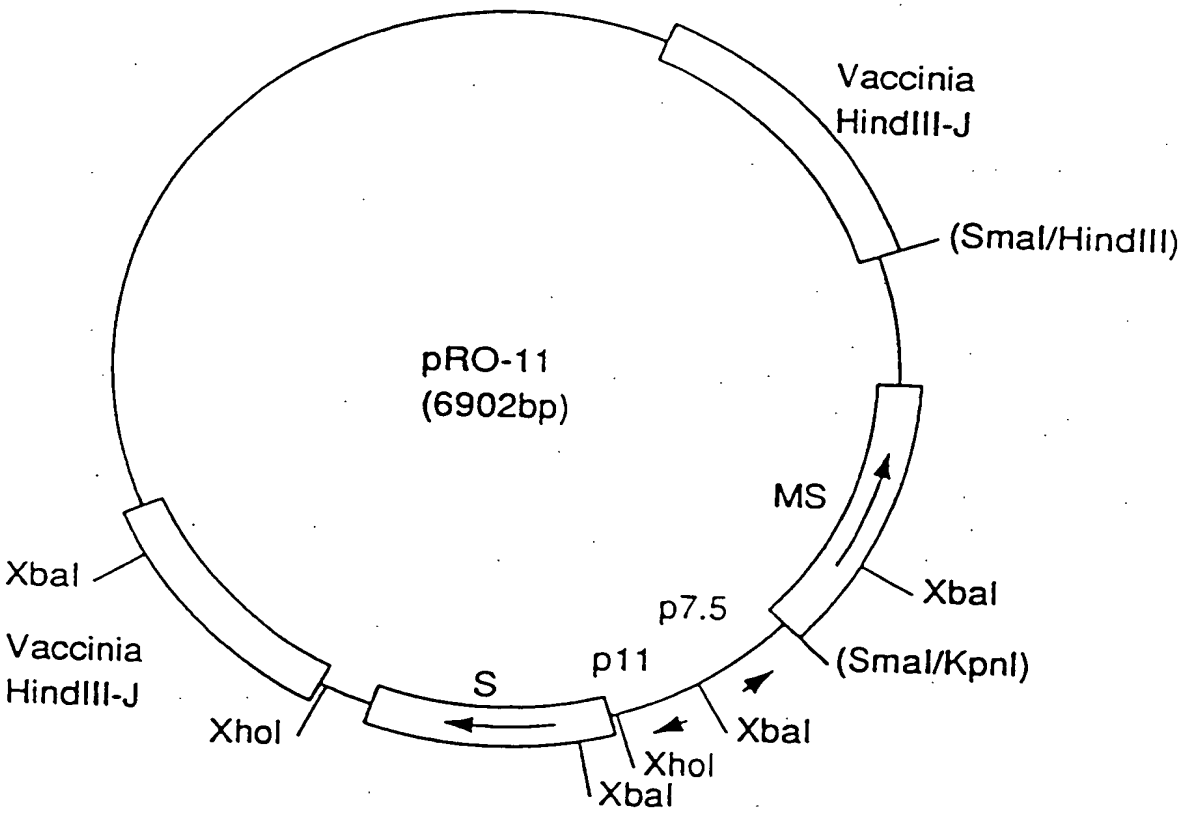

Vaccinia
HindIII-J
(SmaI/HindIII)
pRO-11
(6902bp)
MS
XbaI
p7.5
XbaI
(SmaI/KpnI)
p11
S
XbaI
XhoI
XbaI
Vaccinia
HindIII-J
XhoI

FIG. 55

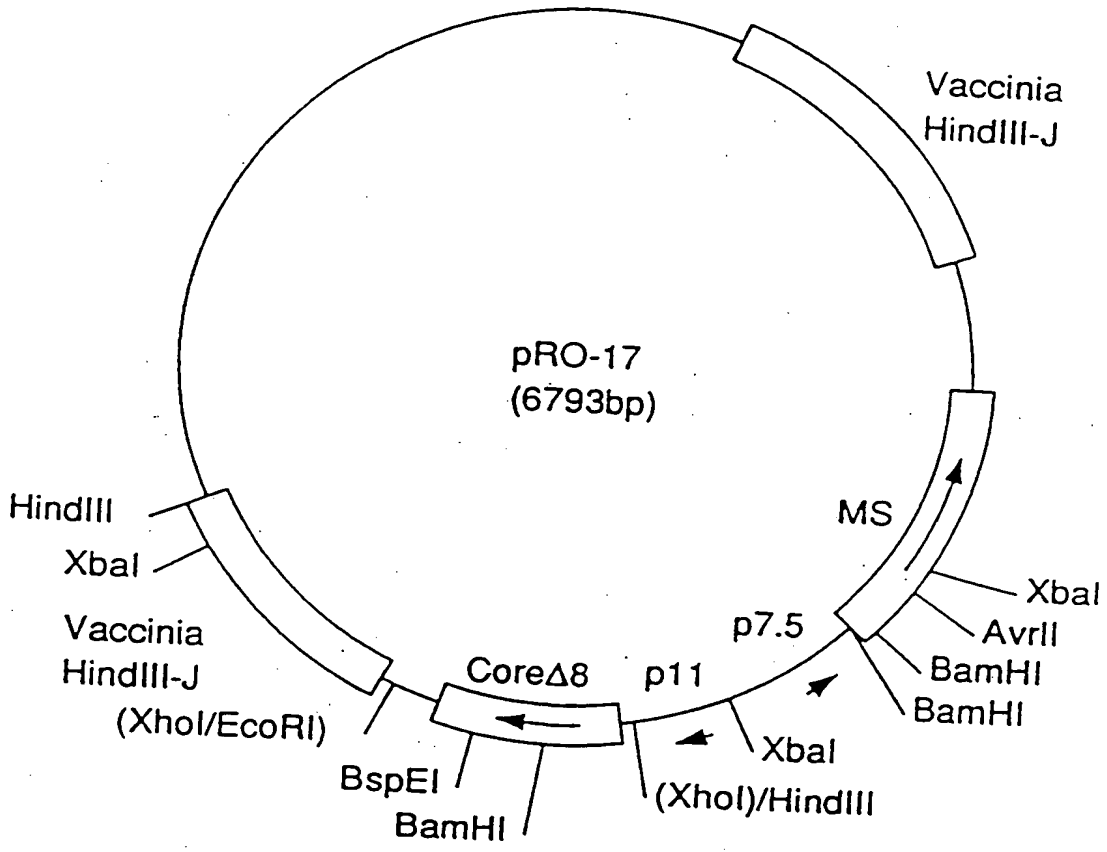

Vaccinia
HindIII-J
pRO-17
(6793bp)
MS
Xbal
AvrII
BamHI
BamHI
p7.5
p11
Xbal
(XhoI)/HindIII
Core∆8
BspEI
BamHI
Vaccinia
HindIII-J
(XhoI/EcoRI)
HindIII
Xbal

FIG. 56

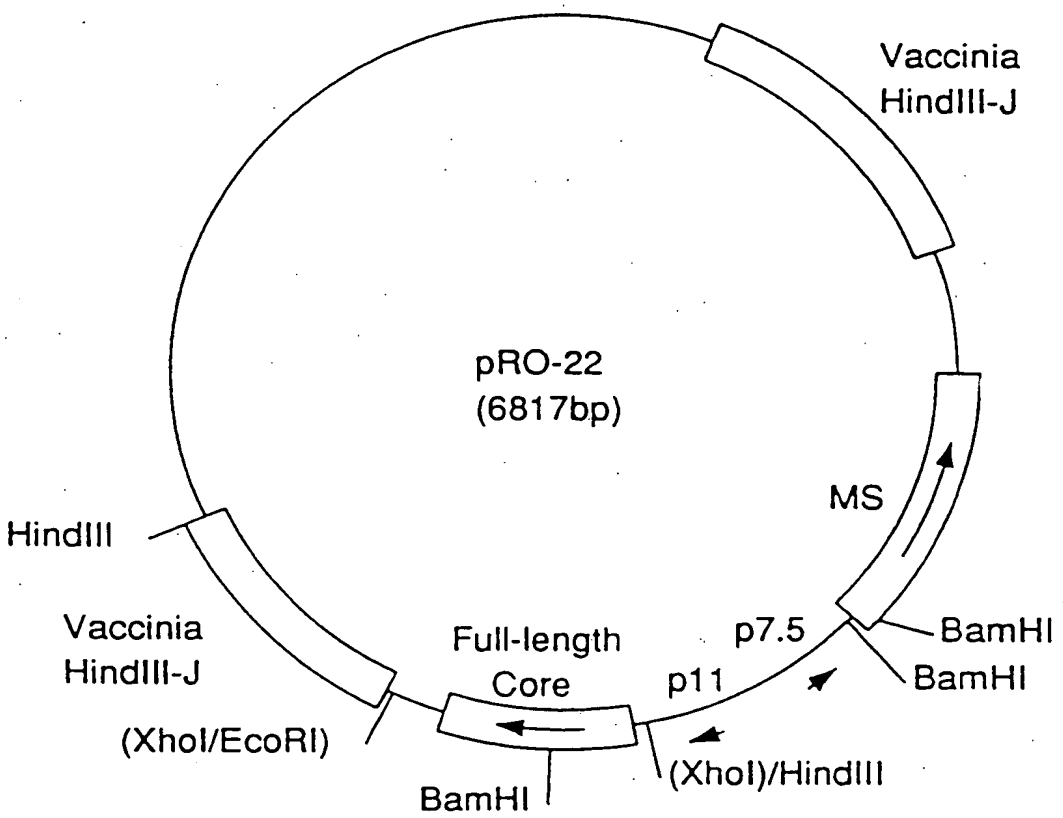

Vaccinia
HindIII-J
pRO-22
(6817bp)
MS
HindIII
Vaccinia
HindIII-J
(XhoI/EcoRI)
Full-length
Core
p11
p7.5
BamHI
BamHI
BamHI
(XhoI)/HindIII

FIG. 57

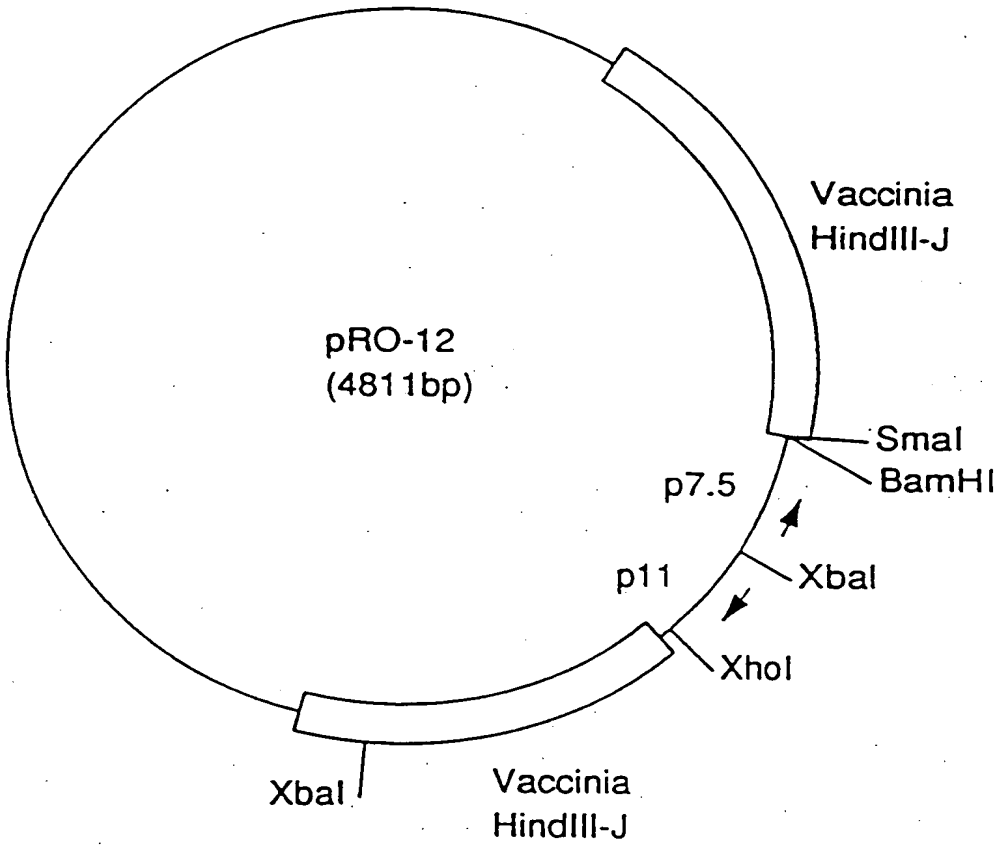

Vaccinia
HindIII-J
pRO-12
(4811bp)
SmaI
BamHI
p7.5
p11
XbaI
XhoI
XbaI
Vaccinia
HindIII-J

FIG. 58

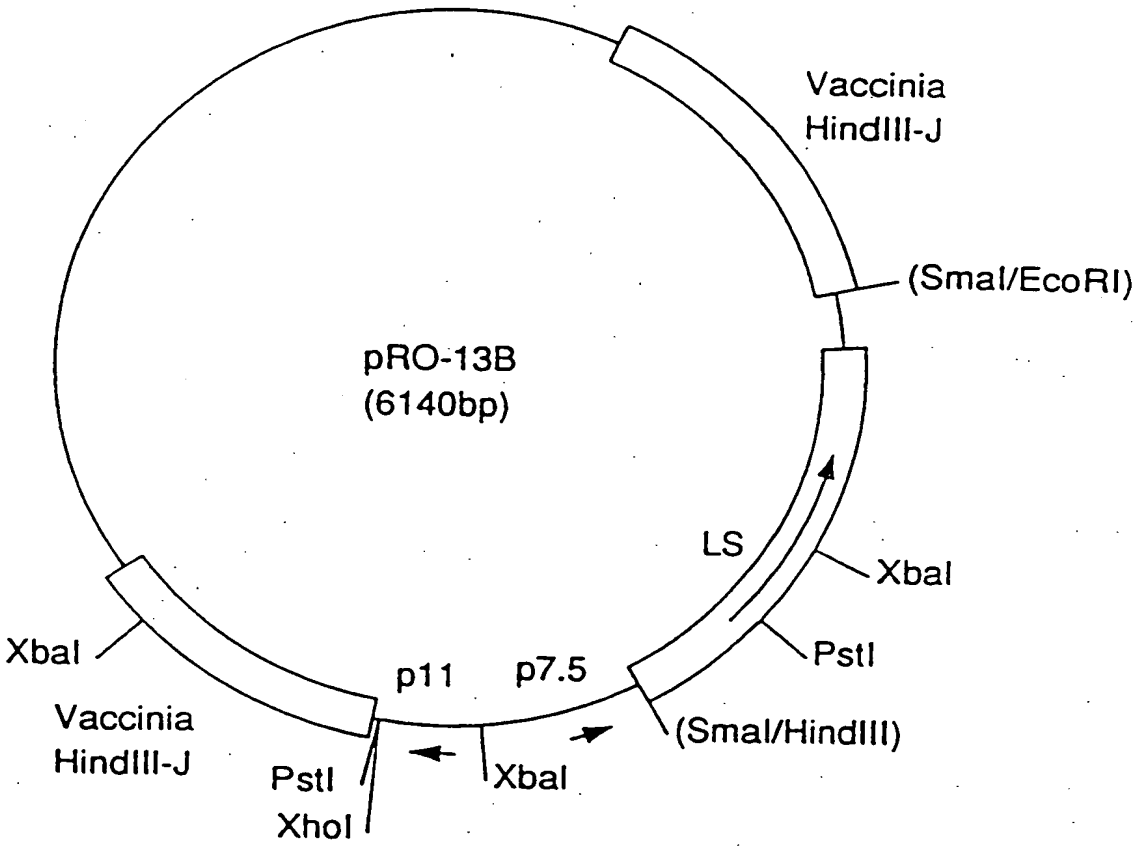

Vaccinia
HindIII-J
(SmaI/EcoRI)
pRO-13B
(6140bp)
LS
XbaI
PstI
(SmaI/HindIII)
p7.5
p11
XbaI
PstI
XhoI
Vaccinia
HindIII-J
XbaI

FIG. 59

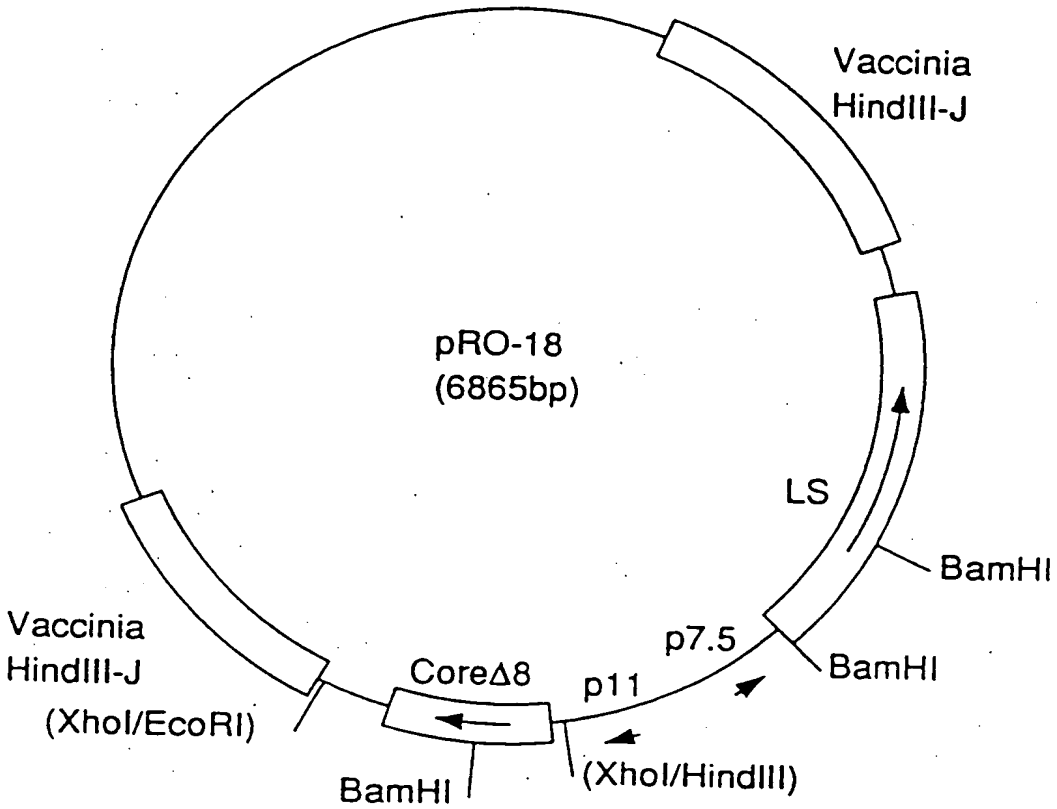

Vaccinia
HindIII-J
pRO-18
(6865bp)
LS
BamHI
p7.5
BamHI
p11
(XhoI/HindIII)
CoreΔ8
BamHI
Vaccinia
HindIII-J
(XhoI/EcoRI)

FIG. 60

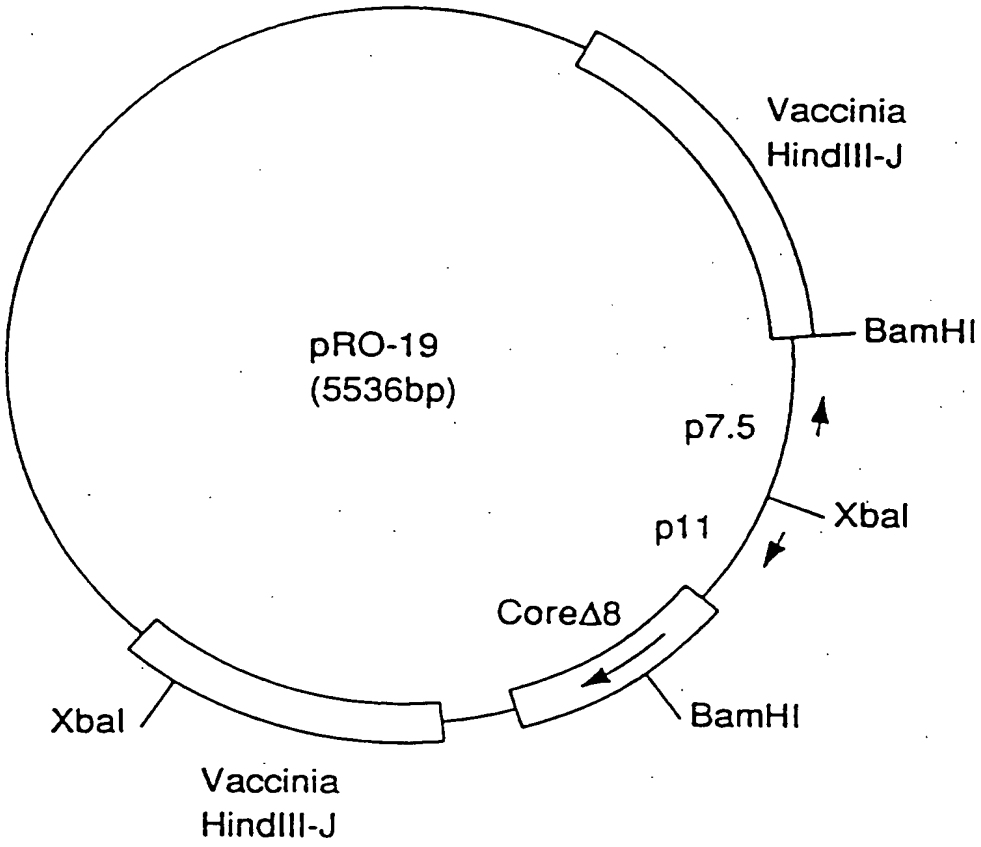

Vaccinia
HindIII-J
BamHI
pRO-19
(5536bp)
p7.5
XbaI
p11
CoreΔ8
BamHI
XbaI
Vaccinia
HindIII-J

FIG. 61

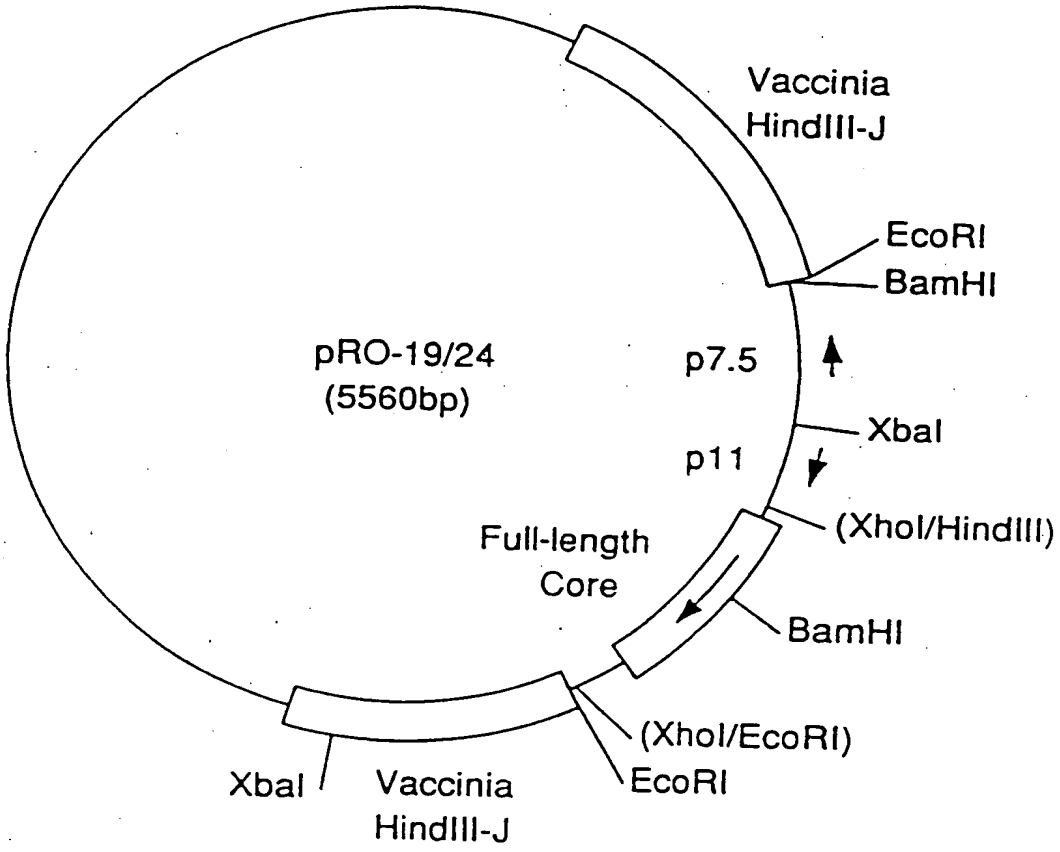

Vaccinia
HindIII-J
EcoRI
BamHI
pRO-19/24
(5560bp)
p7.5
XbaI
p11
(XhoI/HindIII)
Full-length
Core
BamHI
(XhoI/EcoRI)
EcoRI
XbaI
Vaccinia
HindIII-J

FIG. 62

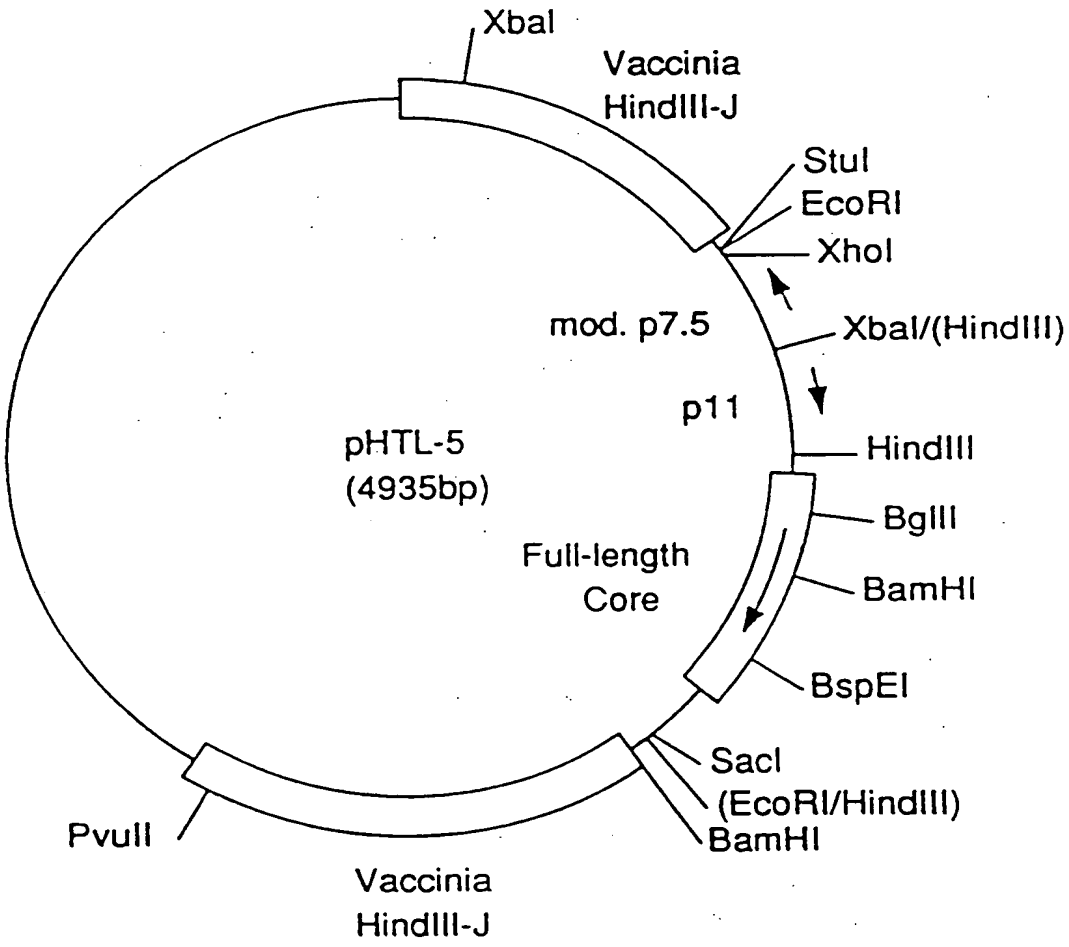

XbaI
Vaccinia
HindIII-J
StuI
EcoRI
XhoI
mod. p7.5
XbaI/(HindIII)
p11
pHTL-5
(4935bp)
HindIII
BglII
Full-length
Core
BamHI
BspEI
SacI
(EcoRI/HindIII)
BamHI
PvuII
Vaccinia
HindIII-J

FIG. 63

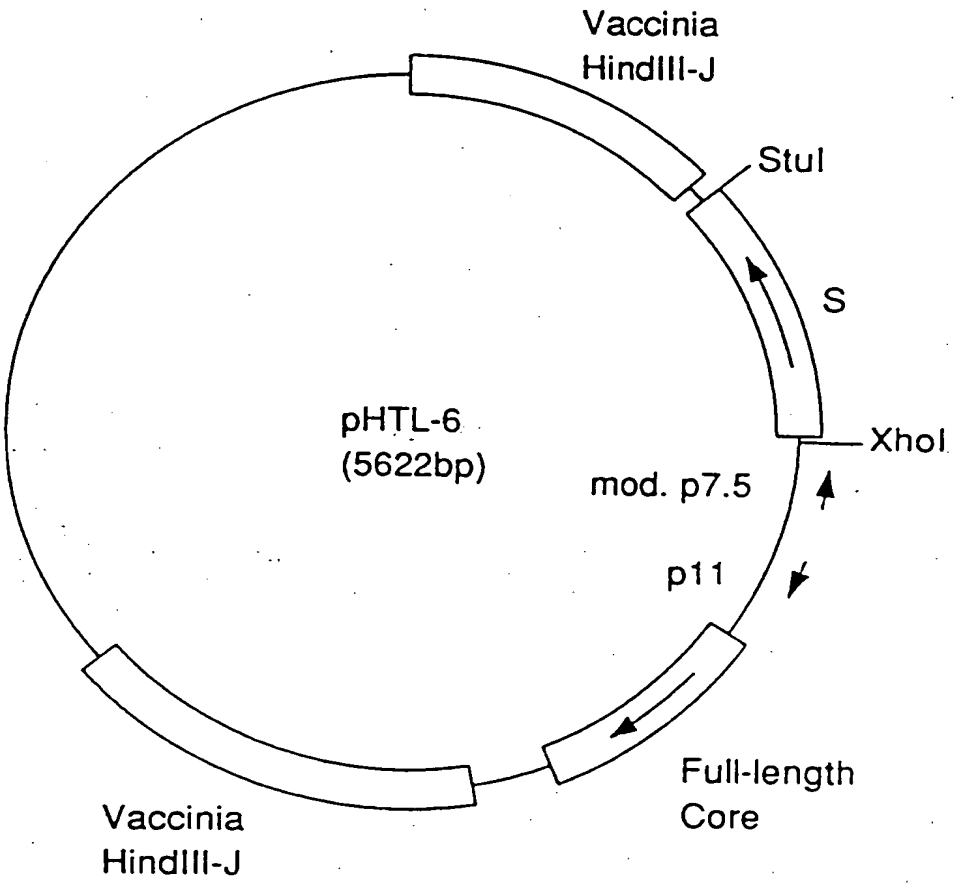

Vaccinia
HindIII-J
StuI
S
XhoI
pHTL-6
(5622bp)
mod. p7.5
p11
Full-length
Core
Vaccinia
HindIII-J

FIG. 64

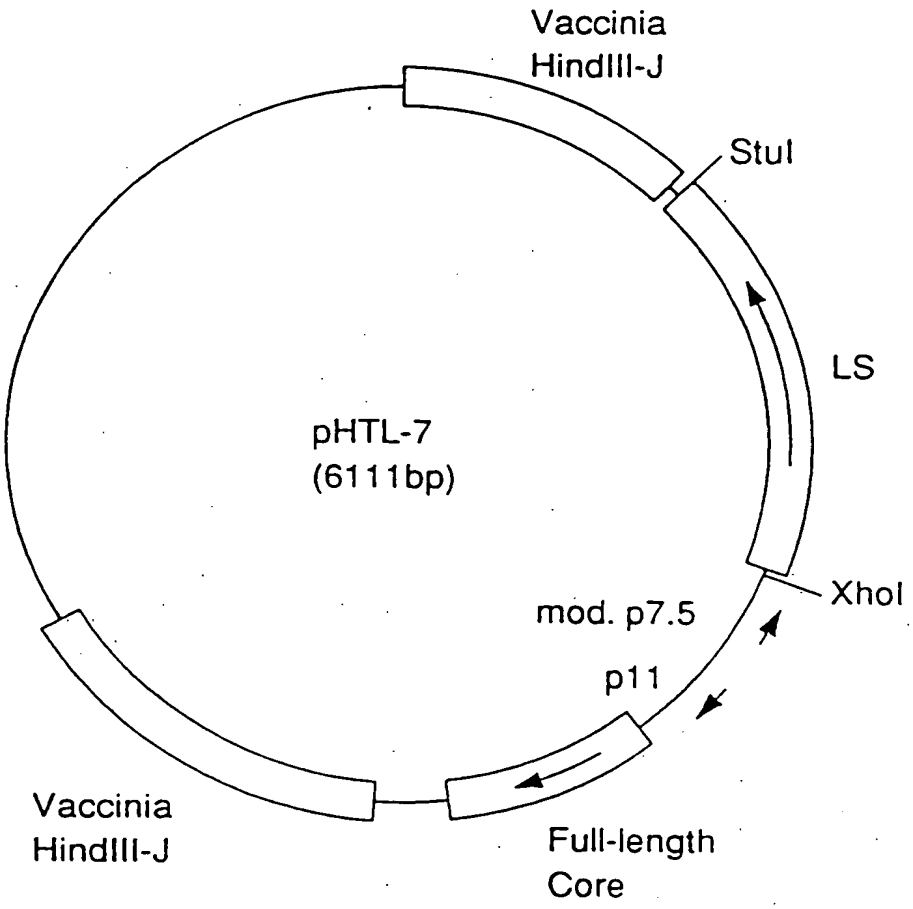

Vaccinia
HindIII-J
StuI
LS
pHTL-7
(6111bp)
XhoI
mod. p7.5
p11
Vaccinia
HindIII-J
Full-length
Core

FIG. 65

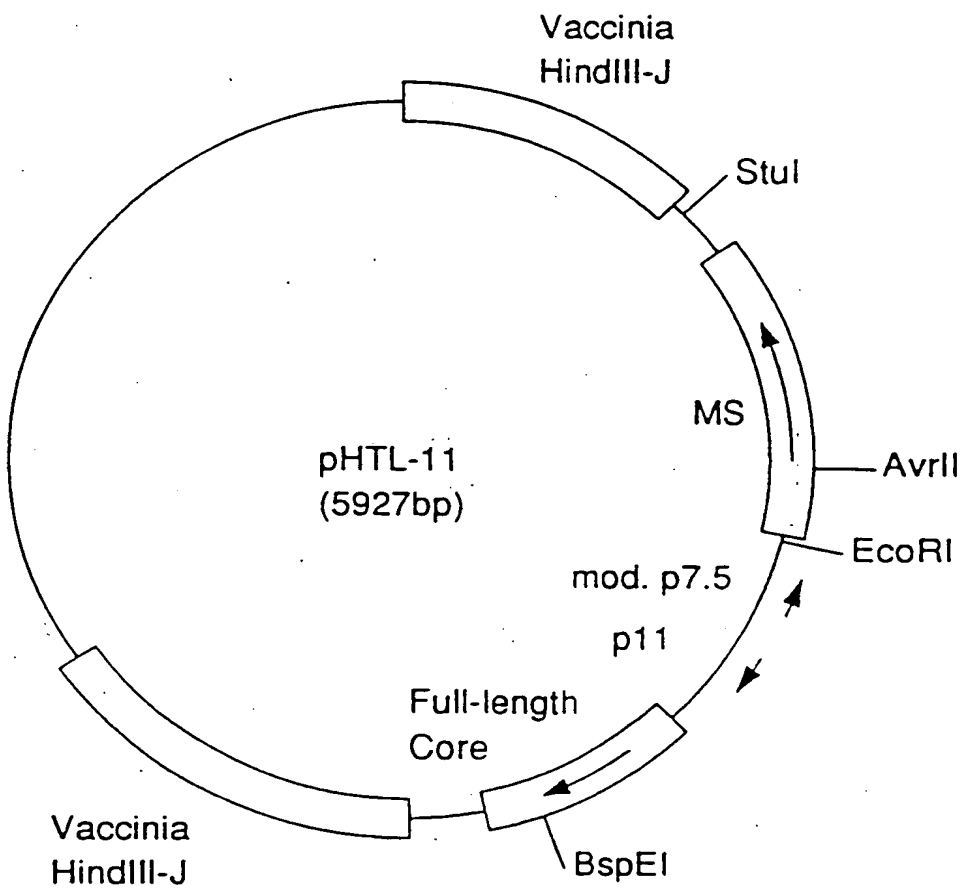

Vaccinia
HindIII-J
StuI
MS
AvrII
EcoRI
mod. p7.5
p11
Full-length
Core
BspEI
Vaccinia
HindIII-J
pHTL-11
(5927bp)

FIG. 66

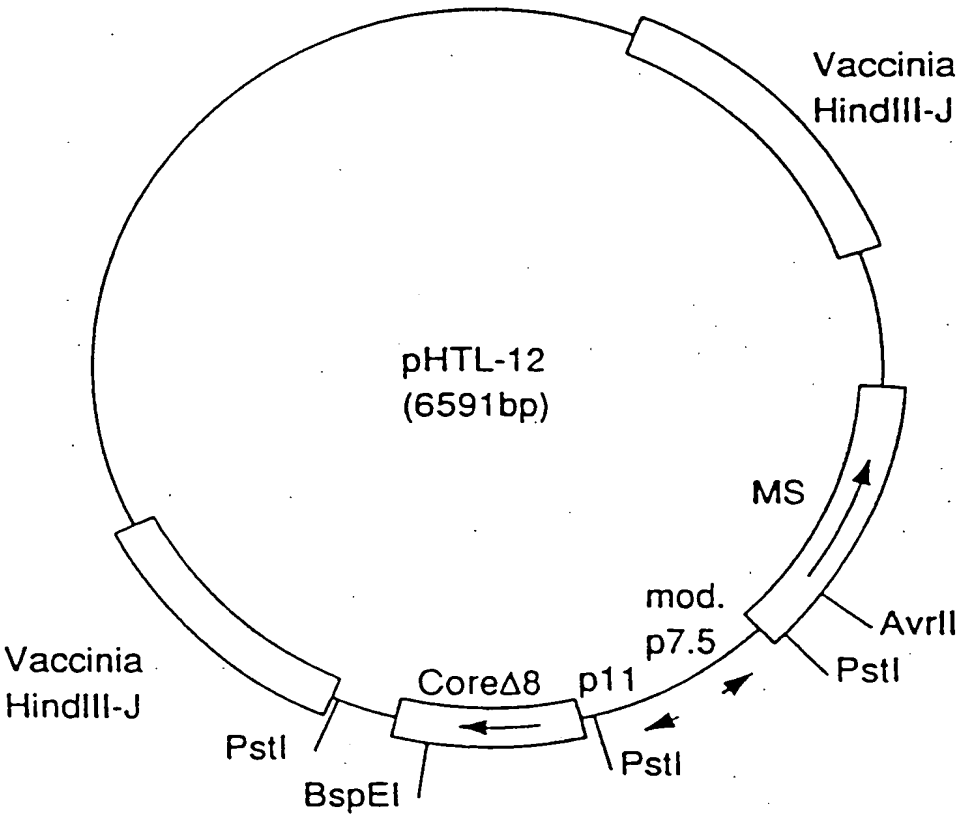

Vaccinia
HindIII-J
pHTL-12
(6591bp)
MS
mod.
p7.5
AvrII
PstI
p11
CoreΔ8
PstI
PstI
BspEI
Vaccinia
HindIII-J

FIG. 67

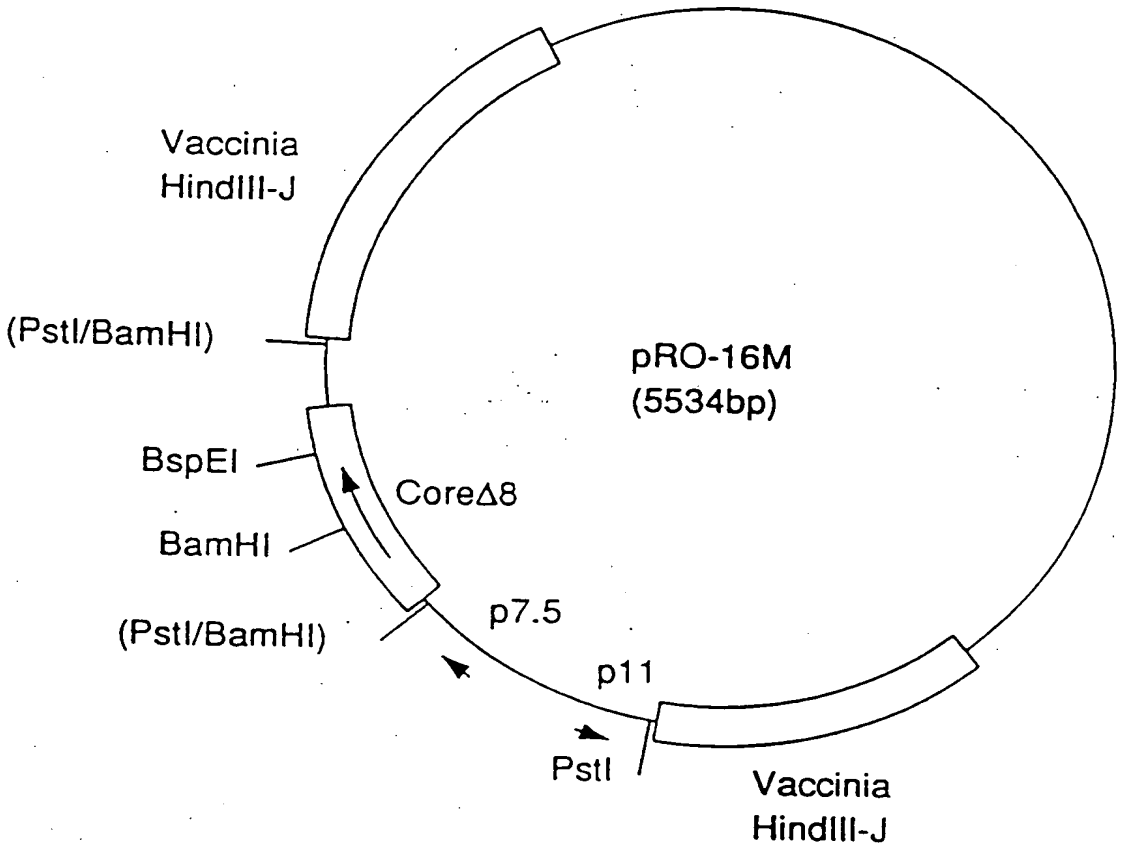

Vaccinia
HindIII-J
(PstI/BamHI)
pRO-16M
(5534bp)
BspEI
CoreΔ8
BamHI
(PstI/BamHI)
p7.5
p11
PstI
Vaccinia
HindIII-J

FIG. 68

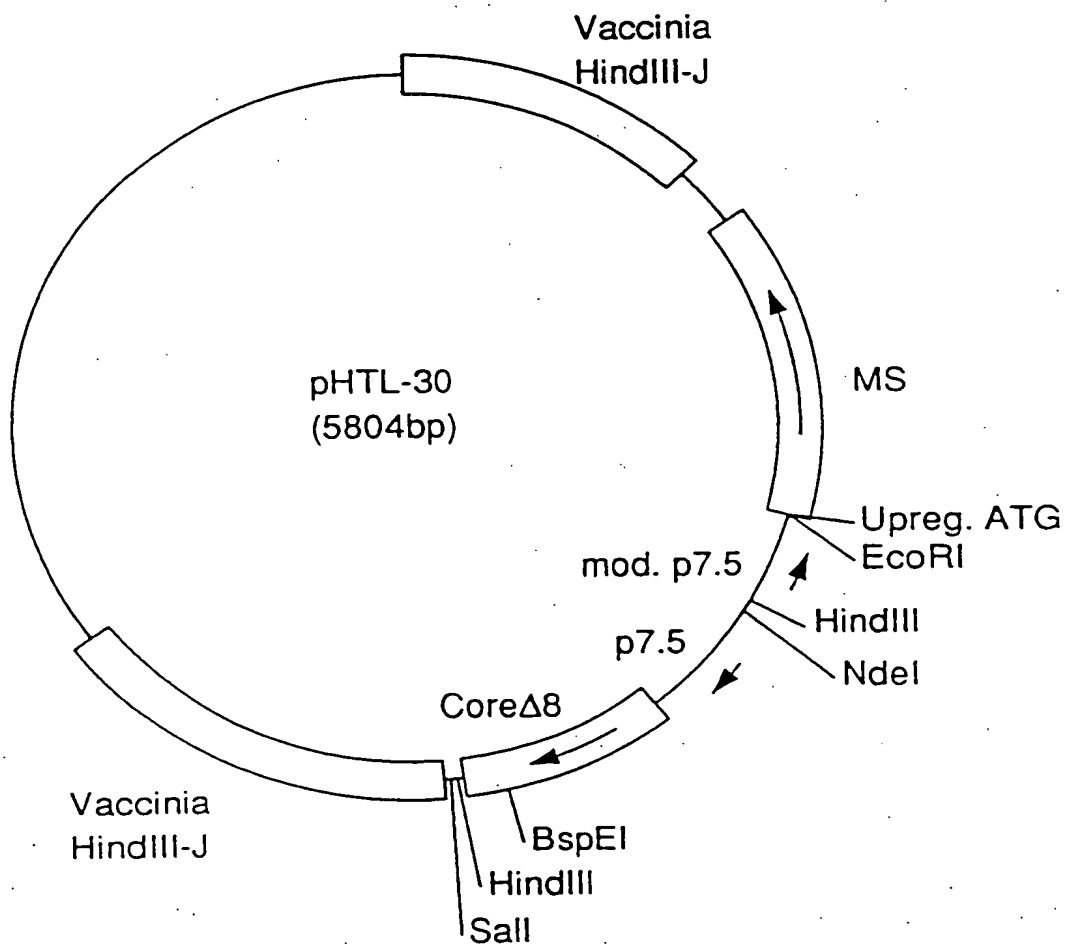

Vaccinia
HindIII-J
pHTL-30
(5804bp)
MS
Upreg. ATG
EcoRI
mod. p7.5
HindIII
p7.5
NdeI
CoreΔ8
Vaccinia
HindIII-J
BspEI
HindIII
SalI

FIG. 69

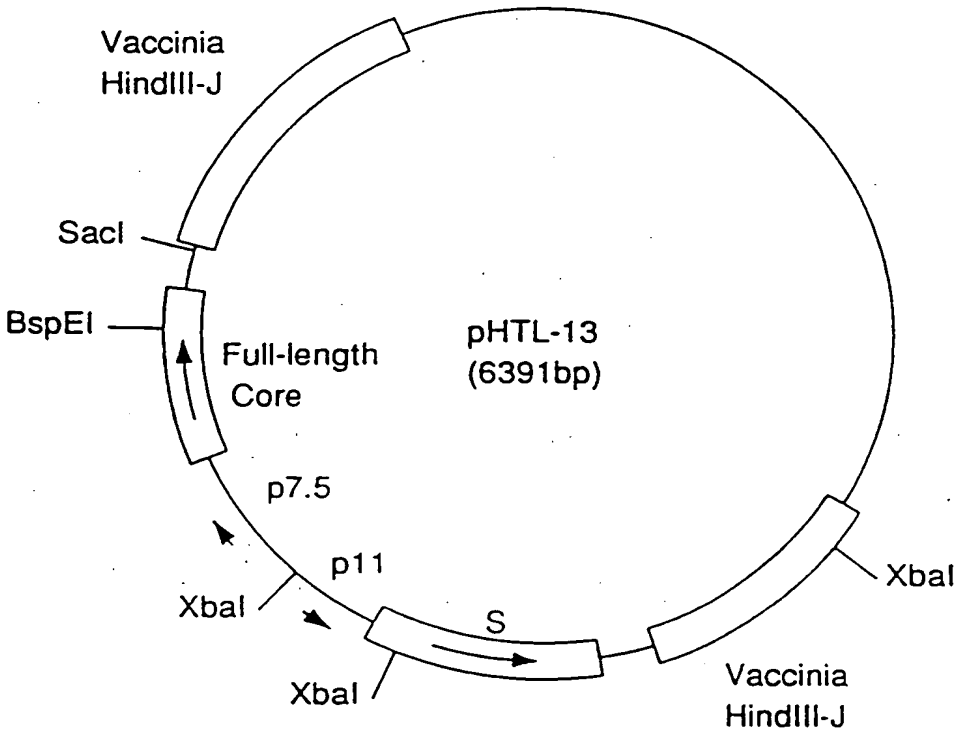

Vaccinia
HindIII-J
SacI
BspEI
Full-length
Core
pHTL-13
(6391bp)
p7.5
p11
XbaI
S
XbaI
XbaI
Vaccinia
HindIII-J

FIG. 70

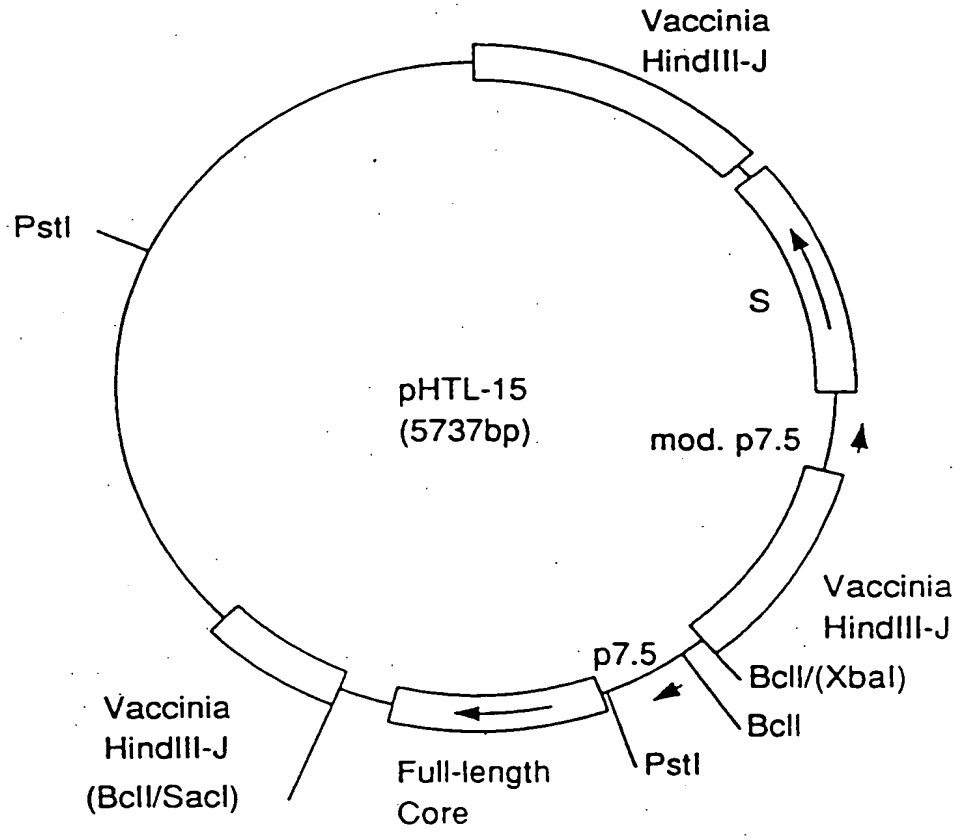

Vaccinia
HindIII-J
PstI
S
pHTL-15
(5737bp)
mod. p7.5
Vaccinia
HindIII-J
p7.5
BclI/(XbaI)
BclI
PstI
Full-length
Core
Vaccinia
HindIII-J
(BclI/SacI)

FIG. 71

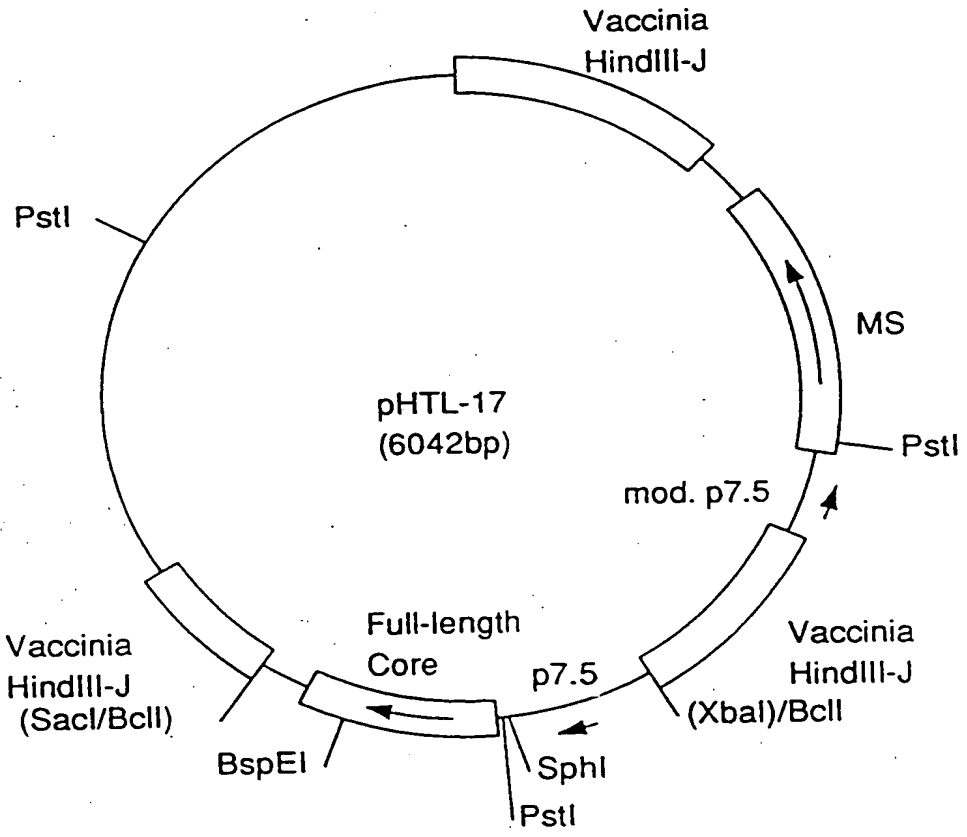

Vaccinia
HindIII-J
PstI
MS
pHTL-17
(6042bp)
PstI
mod. p7.5
Vaccinia
HindIII-J
(XbaI)/BclI
Full-length
Core
p7.5
Vaccinia
HindIII-J
(SacI/BclI)
BspEI
SphI
PstI

FIG. 72

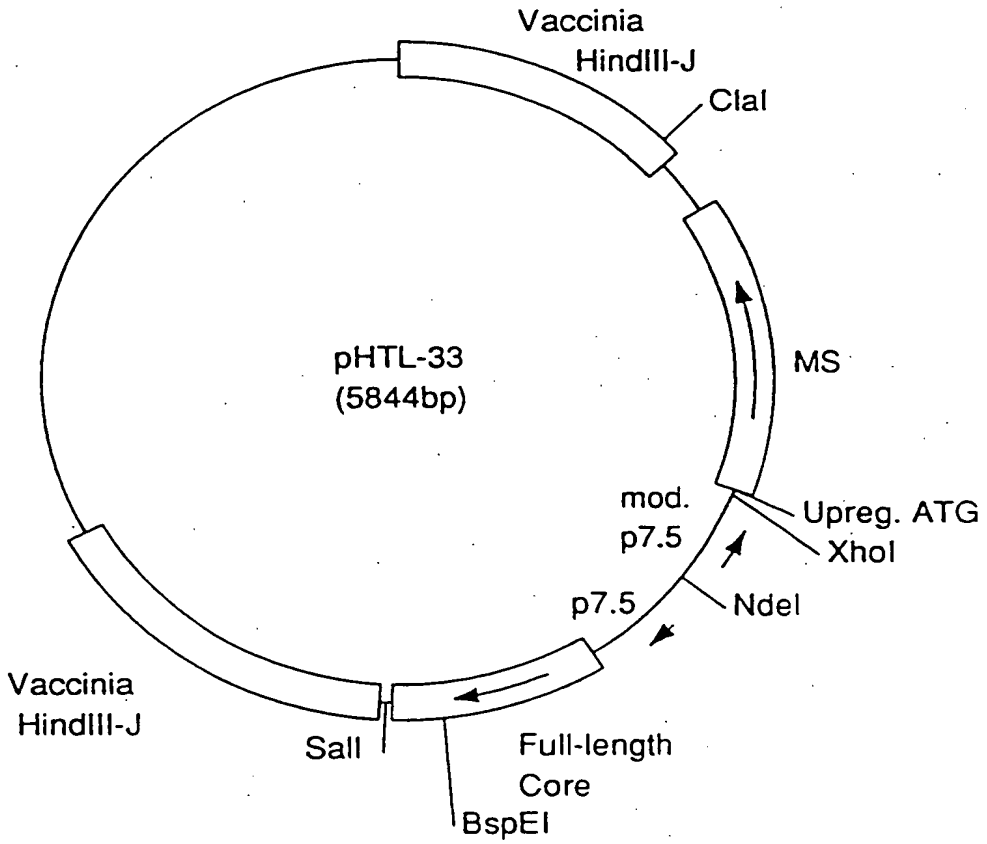

Vaccinia
HindIII-J
ClaI
pHTL-33
(5844bp)
MS
mod.
p7.5
Upreg. ATG
XhoI
p7.5
NdeI
Vaccinia
HindIII-J
SalI
Full-length
Core
BspEI

FIG. 73

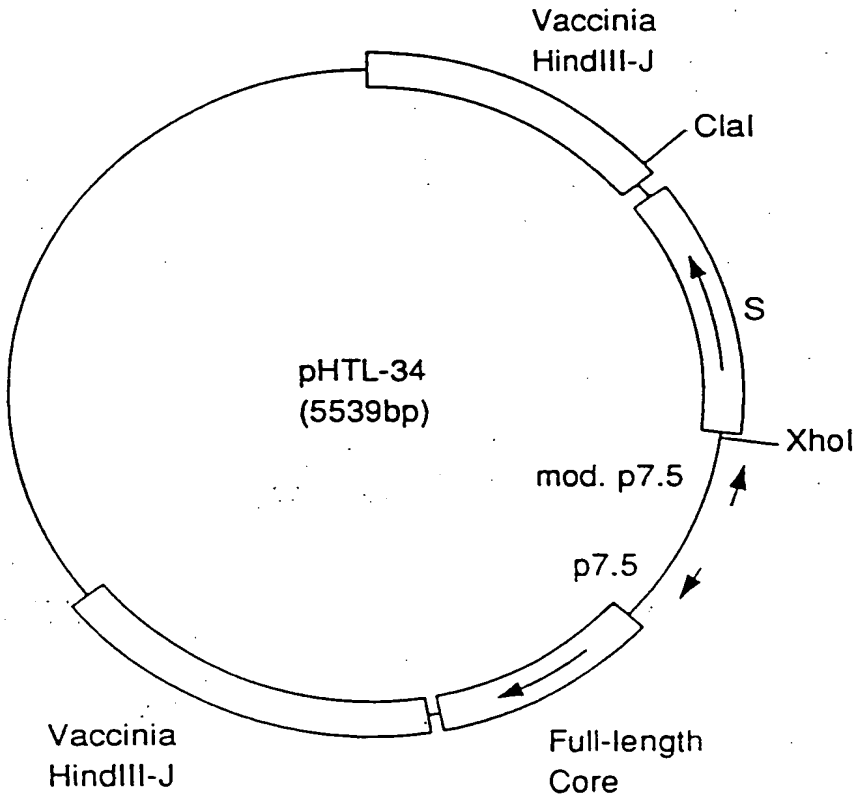

Vaccinia
HindIII-J
ClaI
S
pHTL-34
(5539bp)
XhoI
mod. p7.5
p7.5
Vaccinia
HindIII-J
Full-length
Core

FIG. 74

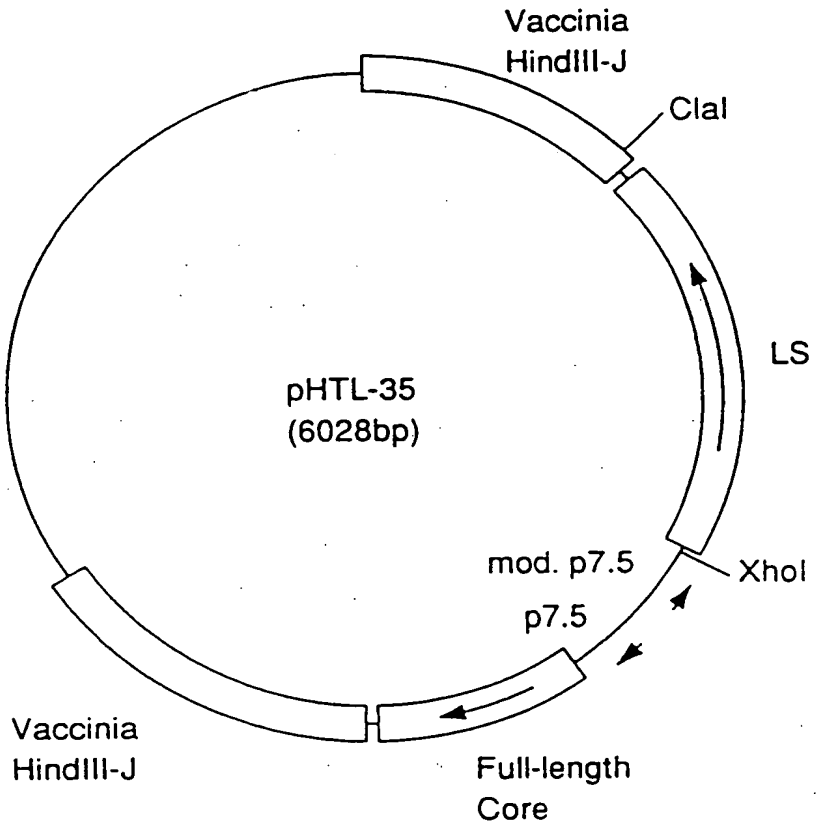

Vaccinia
HindIII-J
ClaI
LS
pHTL-35
(6028bp)
mod. p7.5
XhoI
p7.5
Vaccinia
HindIII-J
Full-length
Core

FIG. 75

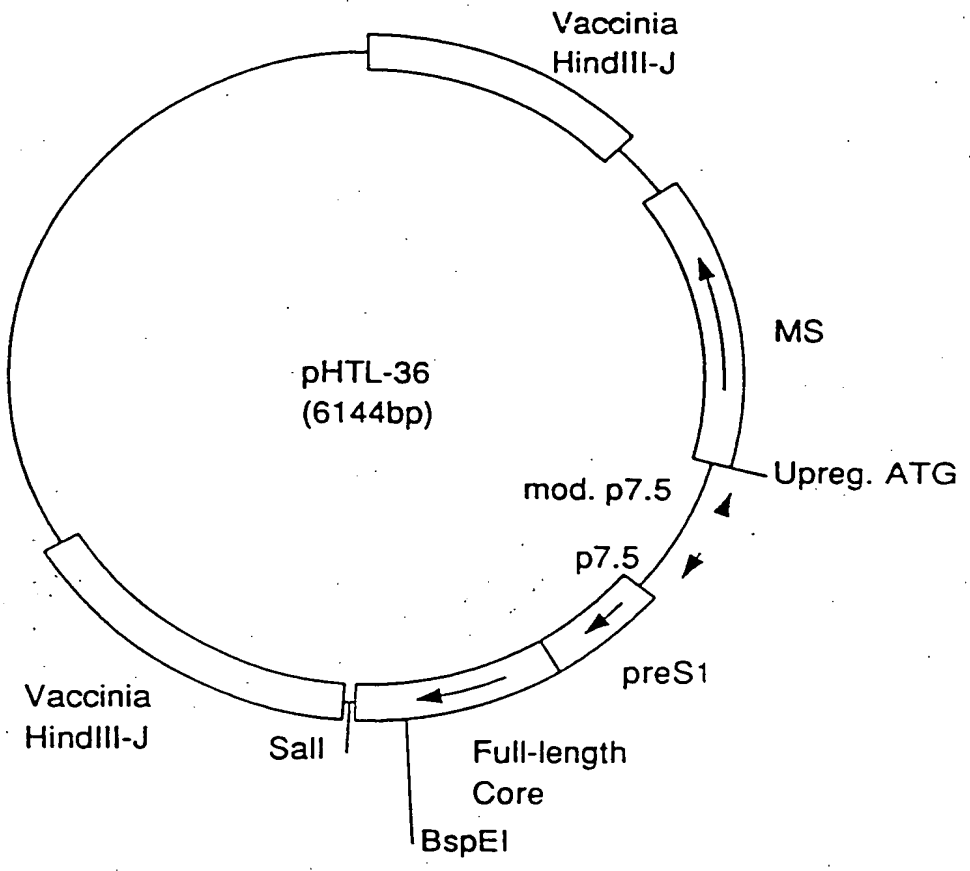

Vaccinia
HindIII-J
MS
pHTL-36
(6144bp)
Upreg. ATG
mod. p7.5
p7.5
preS1
Vaccinia
HindIII-J
SalI
Full-length
Core
BspEI

FIG. 76

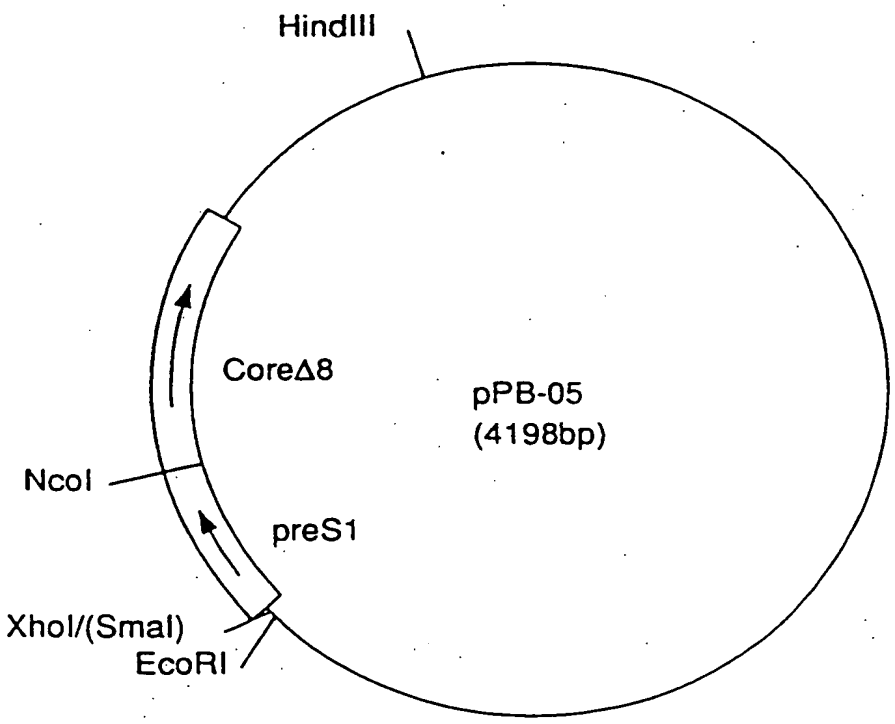


FIG. 77

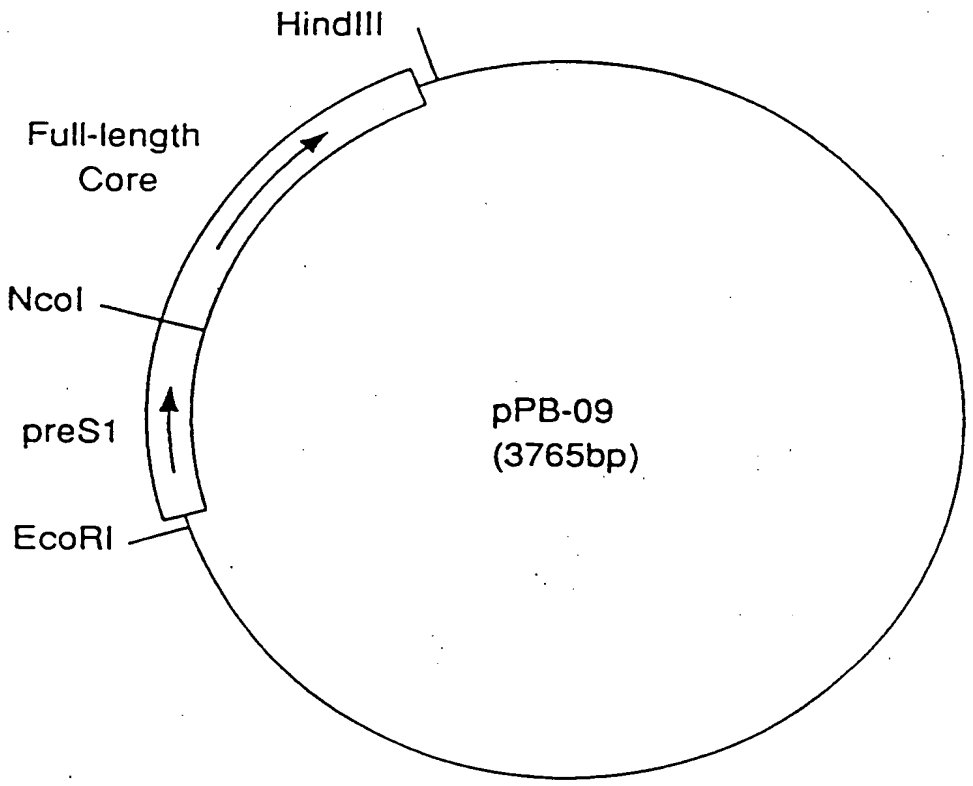


FIG. 78

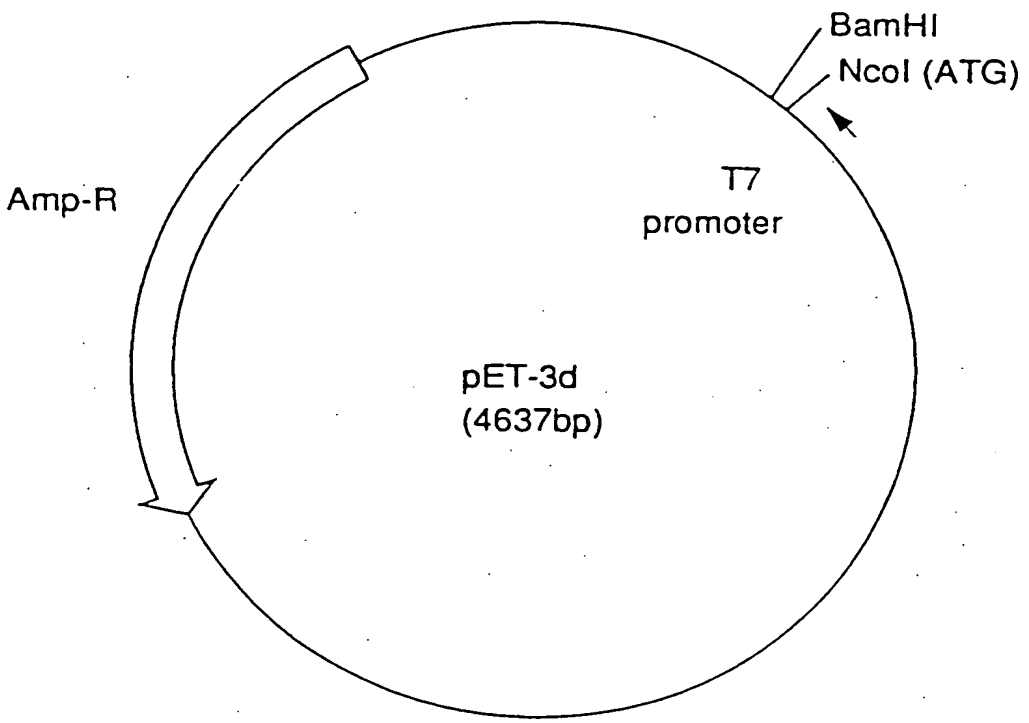


FIG. 79

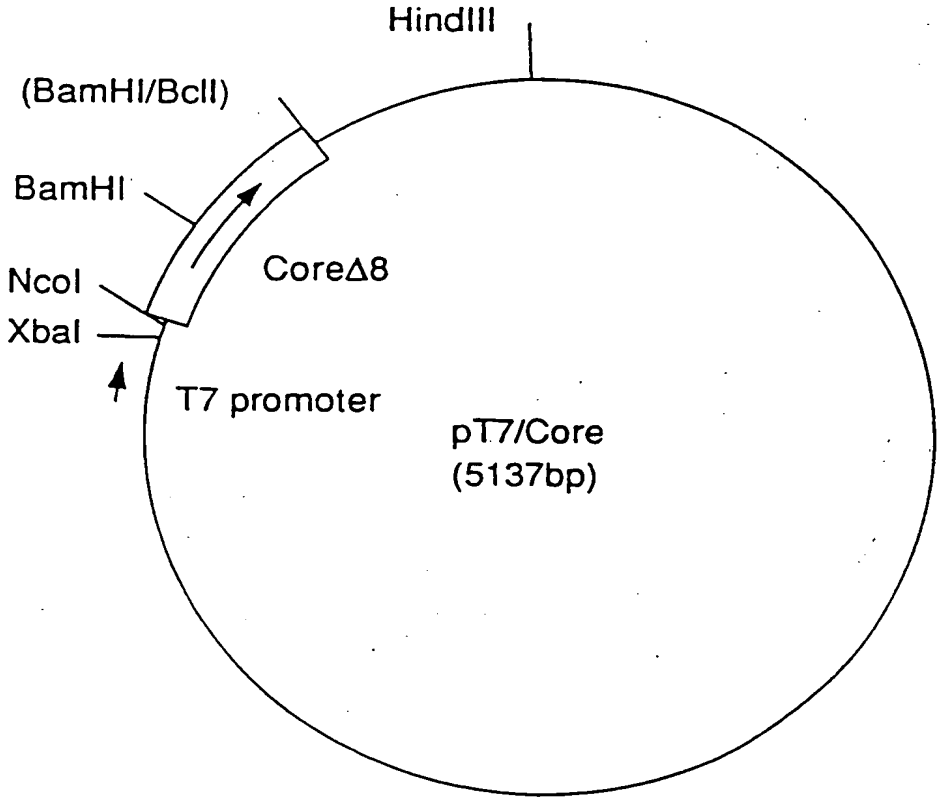
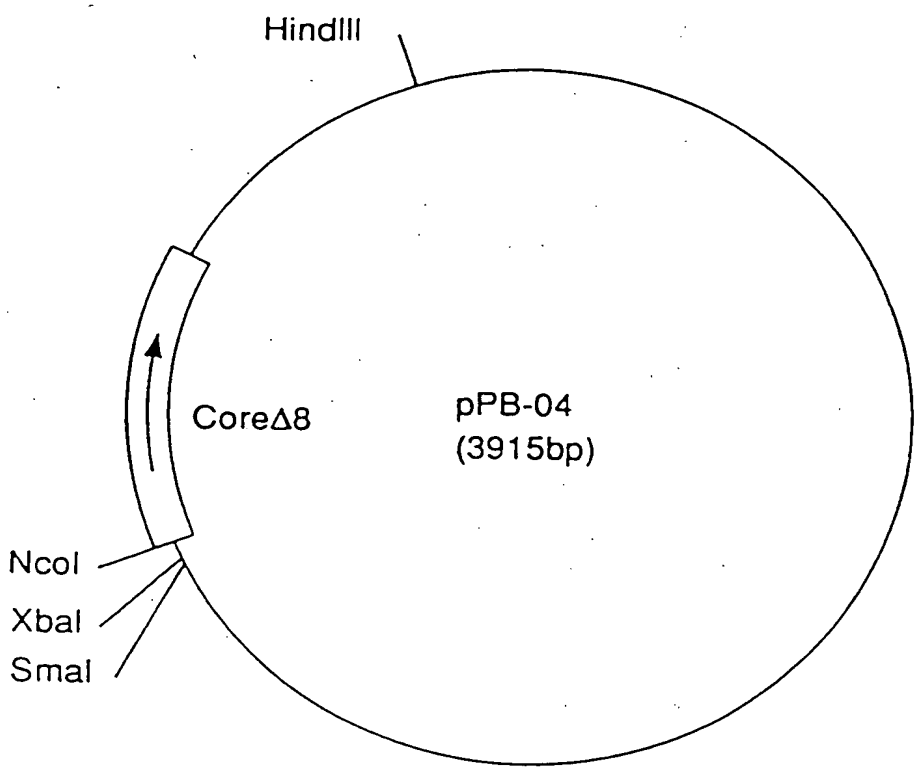

HindIII
(BamHI/BclI)
BamHI
NcoI
XbaI
CoreΔ8
T7 promoter
pT7/Core
(5137bp)

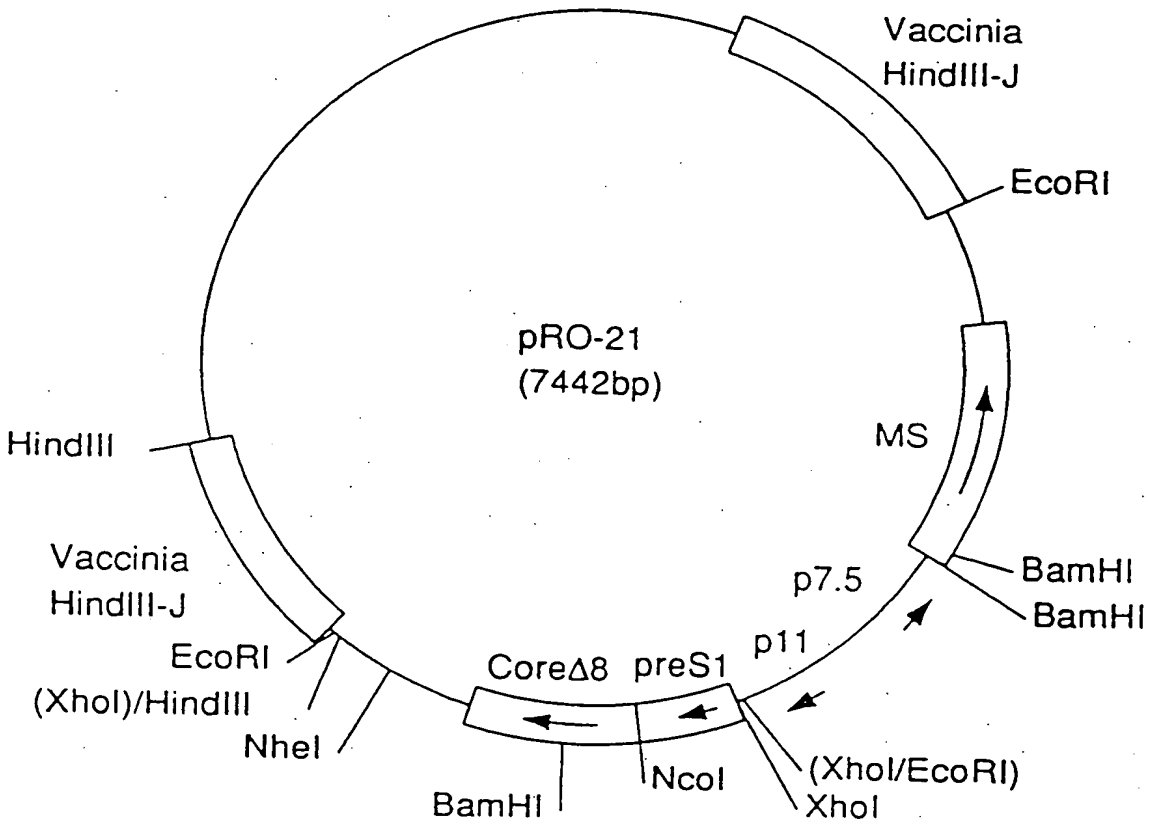
FIG. 80



```
1   /   1                                31   /   11
ATG GGG ACG AAT CTT TCT GTT CCC AAt CCT CTG GGA TTC tTT CCC GAT CAT CAG TTG CAC
met gly thr asn leu ser val pro asn pro leu gly phe phe pro asp his gln leu asp
61   /   21                              91   /   31
CCT GCA TTC GGA GCC AAC TCA Aac AAT CCA GAT TGG GAC TTC AAC CCC ATC AAG GAC CAC
pro ala phe gly ala asn ser asn asn pro asp trp asp phe asn pro ile lys asp his
121   /   41                            151   /   51
TGG CCA GCA GCC AAC CAG GTA GGA GTG GGA GCA TTC GGG CCA GGG tTC ACC CCT CCA CAC
trp pro ala ala asn gln val gly val gly ala phe gly pro gly phe thr pro pro his
181   /   61                            211   /   71
GGC GGT gTT TTG GGG TGG AGC CCT CAG GCT CAG GGC ATA TTG ACC ACA GTG TCA ACA ATT
gly gly val leu gly trp ser pro gln ala gln gly ile leu thr thr val ser thr ile
241   /   81                            271   /   91
CCT CCT CCT GCC TCC ACC AAT CGG CAG TCA GGA AGG CAG CCT ACT CCC ATC TCT CCA CCT
pro pro pro ala ser thr asn arg gln ser gly arg gln pro thr pro ile ser pro pro
301   /   101                           331   /   111
CTA AGA GAC AGT CAT CCT CAG GCC ATG GAC ATT GAC CCT TAT AAA GAA TTT GGA GCT ACT
leu arg asp ser his pro gln ala met asp ile asp pro tyr lys glu phe gly ala thr
361   /   121                           391   /   131
GTG GAG TTA CTC TCG TTT TTG CCT TCT GAC TTC TTT CCT TCC GTC AGA GAT CTC CTA GAC
val glu leu leu ser phe leu pro ser asp phe phe pro ser val arg asp leu leu asp
421   /   141                           451   /   151
ACC GCC TCA GCT CTG TAT CGG GAA GCC TTA GAG TCT CCT GAG CAT TGC TCA CCT CAC CAC
thr ala ser ala leu tyr arg glu ala leu glu ser pro glu his cys ser pro his his
481   /   161                           511   /   171
ACC GCA CTC AGG CAA GCC ATT CTC TGC TGG GGG GAA TTG ATG ACT CTA GCT ACC TGG GTG
thr ala leu arg gln ala ile leu cys trp gly glu leu met thr leu ala thr trp val
541   /   181                           571   /   191
GGT AAT AAT TTG GAG GAT CCA GCA TCA AGG GAT CTA GTA GTC AAT TAT GTT AAT ACT AAC
gly asn asn leu glu asp pro ala ser arg asp leu val val asn tyr val asn thr asn
601   /   201                           631   /   211
ATG GGT TTA AAA ATT AGG CAA CTA TTG TGG TTT CAT ATA TCT TGC CTT ACT TTT GGA AGA
met gly leu lys ile arg gln leu leu trp phe his ile ser cys leu thr phe gly arg
661   /   221                           691   /   231
GAG ACT GTA CTT GAA TAT TTG GTA TCT TTC GGA GTG TGG ATT CGC ACT CCT CCA GCC TAT
glu thr val leu glu tyr leu val ser phe gly val trp ile arg thr pro pro ala tyr
721   /   241                           751   /   251
AGA CCA CCA AAT GCC CCT ATC TTA TCA ACA CTT CCG GAA ACT ACT GTT GTT AGA CGA CGG
arg pro pro asn ala pro ile leu ser thr leu pro glu thr thr val val arg arg arg
781   /   261                           811   /   271
GAC CGA GGC AGG TCC CCT AGA AGA AGA ACT CCC TCG CCT CGC AGA CGC AGA TCT CAA TCT
asp arg gly arg ser pro arg arg arg thr pro ser pro arg arg arg arg ser gln ser
841   /   281
CGG GAA TCT CAA TGT TAG
arg glu ser gln cys AMB
```

FIG.81

FIG. 82


Vaccinia
HindIII-J
EcoRI
pRO-21
(7442bp)
MS
BamHI
BamHI
p7.5
p11
Core∆8
preS1
NcoI
(XhoI/EcoRI)
XhoI
BamHI
NheI
(XhoI)/HindIII
EcoRI
Vaccinia
HindIII-J
HindIII

```
1   /   1                                   31  /  11
Atg ggg acg aat ctt tct gtt ccc aat cct ctg gga ttc ttt ccc gat cat cag ttg gac
met gly thr asn leu ser val pro asn pro leu gly phe phe pro asp his gln leu asp
61  /  21                                   91  /  31
cct gca ttc gga gcc aac tca aac aat cca gat tgg gac ttc aac ccc atc aag gac cac
pro ala phe gly ala asn ser asn asn pro asp trp asp phe asn pro ile lys asp his
121 /  41                                   151 /  51
tgg cca gca gcc aac cag gta gga gtg gga gca ttc ggg cca ggg ttc acc cct cca cac
trp pro ala ala asn gln val gly val gly ala phe gly pro gly phe thr pro pro his
181 /  61                                   211 /  71
ggc ggt gtt ttg ggg tgg agc cct cag gct cag ggc ata ttg acc aca gtg tca aca att
gly gly val leu gly trp ser pro gln ala gln gly ile leu thr thr val ser thr ile
241 /  81                                   271 /  91
cct cct cct gcc tcc acc aat cgg cag tca gga agg cag cct act ccc atc tct cca cct
pro pro pro ala ser thr asn arg gln ser gly arg gln pro thr pro ile ser pro pro
301 / 101                                   331 / 111
cta aga gac agt cat cct cag gcc ATG GAC ATT GAC CCT TAT AAA GAA TTT GGA GCT ACT
leu arg asp ser his pro gln ala met asp ile asp pro tyr lys glu phe gly ala thr
361 / 121                                   391 / 131
GTG GAG TTA CTC TCG TTT TTG CCT TCT GAC TTC TTT CCT TCC GTC AGA GAT CTC CTA GAC
val glu leu leu ser phe leu pro ser asp phe phe pro ser val arg asp leu leu asp
421 / 141                                   451 / 151
ACC GCC TCA GCT CTG TAT CGG GAA GCC TTA GAG TCT CCT GAG CAT TGC TCA CCT CAC CAC
thr ala ser ala leu tyr arg glu ala leu glu ser pro glu his cys ser pro his his
481 / 161                                   511 / 171
ACC GCA CTC AGG CAA GCC ATT CTC TGC TGG GGG GAA TTG ATG ACT CTA GCT ACC TGG GTG
thr ala leu arg gln ala ile leu cys trp gly glu leu met thr leu ala thr trp val
541 / 181                                   571 / 191
GGT AAT AAT TTG GAG GAT CCA GCA TCA AGG GAT CTA GTA GTC AAT TAT GTT AAT ACT AAC
gly asn asn leu glu asp pro ala ser arg asp leu val val asn tyr val asn thr asn
601 / 201                                   631 / 211
ATG GGT TTA AAA ATT AGG CAA CTA TTG TGG TTT CAT ATA TCT TGC CTT ACT TTT GGA AGA
met gly leu lys ile arg gln leu leu trp phe his ile ser cys leu thr phe gly arg
661 / 221                                   691 / 231
GAG ACT GTA CTT GAA TAT TTG GTA TCT TTC GGA GTG TGG ATT CGC ACT CCT CCA GCC TAT
glu thr val leu glu tyr leu val ser phe gly val trp ile arg thr pro pro ala tyr
721 / 241                                   751 / 251
AGA CCA CCA AAT GCC CCT ATC TTA TCA ACA CTT CCG GAA ACT ACT GTT GTT AGA CGA CGG
arg pro pro asn ala pro ile leu ser thr leu pro glu thr thr val val arg arg arg
781 / 261                                   811 / 271
GAC CGA GGC AGG TCC CCT AGA AGA AGA ACT CCC TCG CCT CGC AGA CGC AGA tcc caa tcg
asp arg gly arg ser pro arg arg arg thr pro ser pro arg arg arg arg ser gln ser
841 / 281                                   871 / 291
ccg cgt cgc aga cga tct CAA TCT CGG GAA TCT CAA TGT TAG
pro arg arg arg arg ser gln ser arg glu ser gln cys AMB
```

FIG.83

FIG. 84

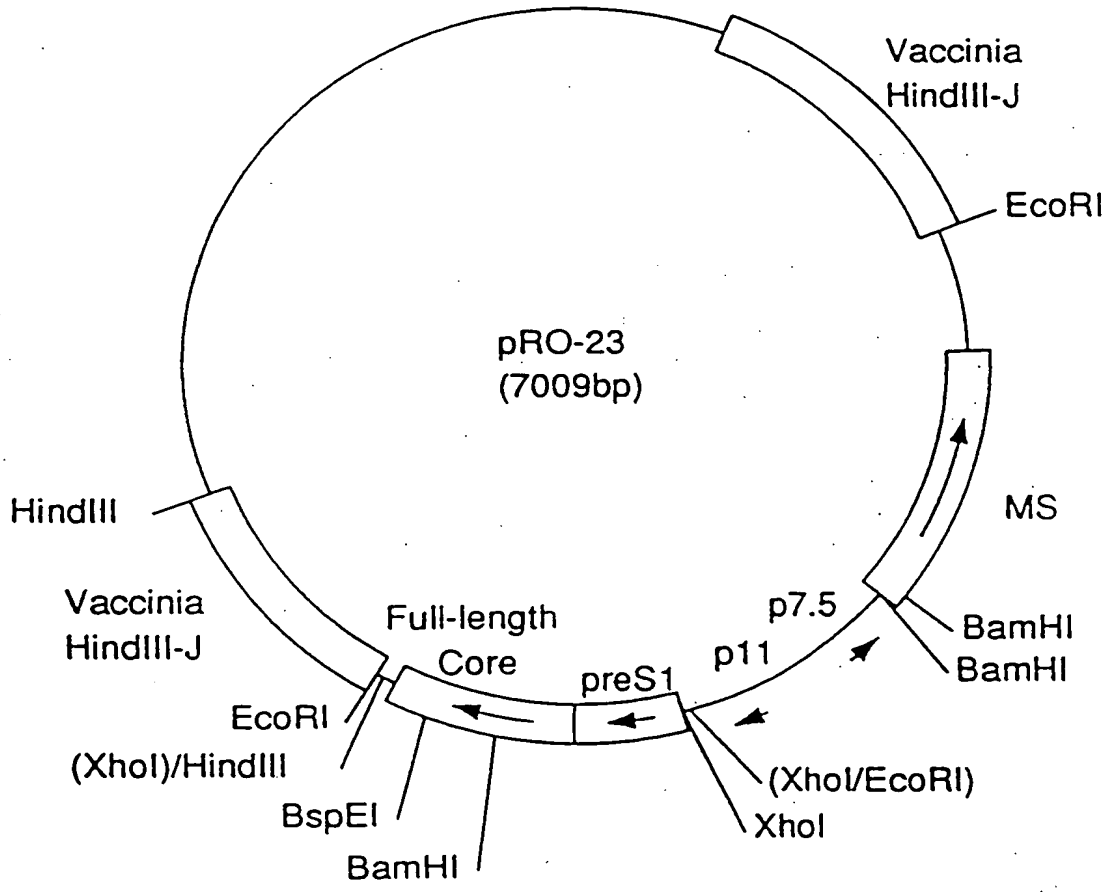

Vaccinia
HindIII-J
EcoRI
pRO-23
(7009bp)
MS
BamHI
BamHI
p7.5
p11
preS1
Full-length
Core
Vaccinia
HindIII-J
HindIII
EcoRI
(XhoI)/HindIII
BspEI
BamHI
(XhoI/EcoRI)
XhoI

FIG. 85

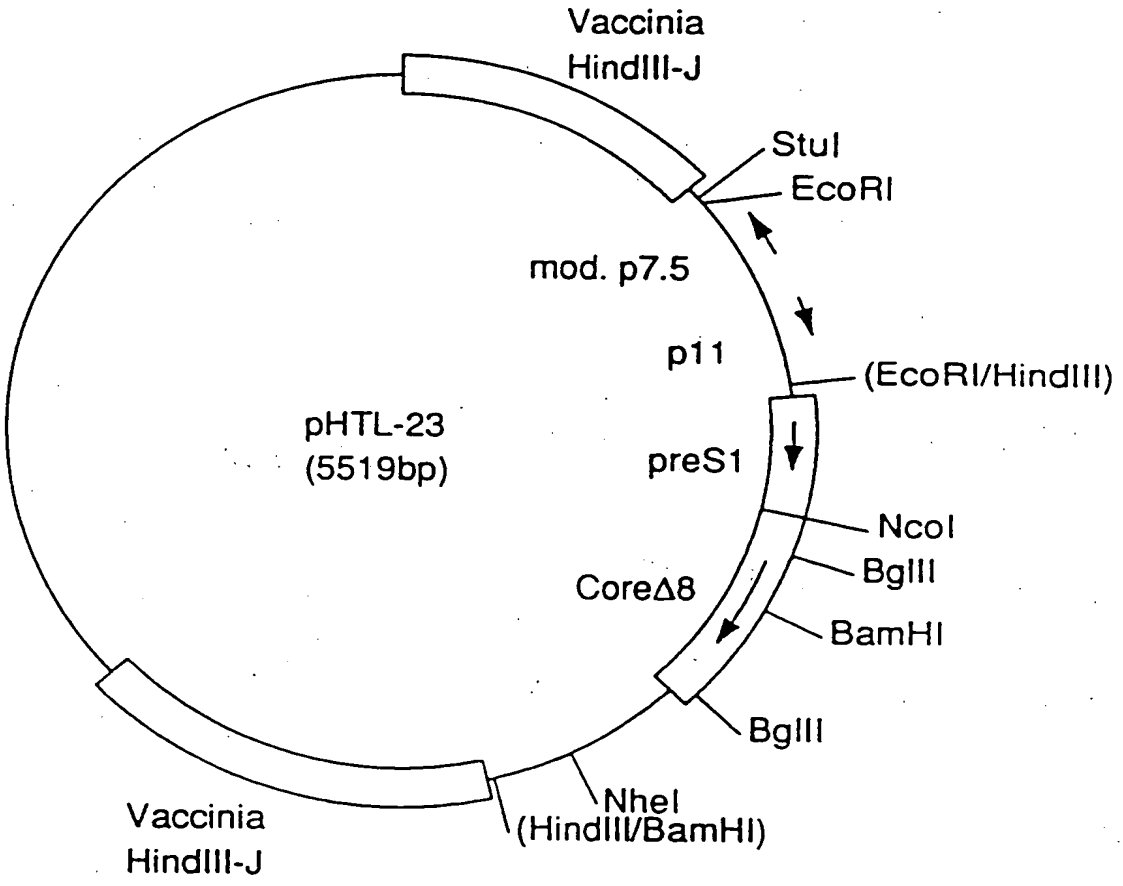

Vaccinia
HindIII-J
StuI
EcoRI
mod. p7.5
p11
(EcoRI/HindIII)
pHTL-23
(5519bp)
preS1
NcoI
BglII
CoreΔ8
BamHI
BglII
NheI
(HindIII/BamHI)
Vaccinia
HindIII-J

FIG. 86

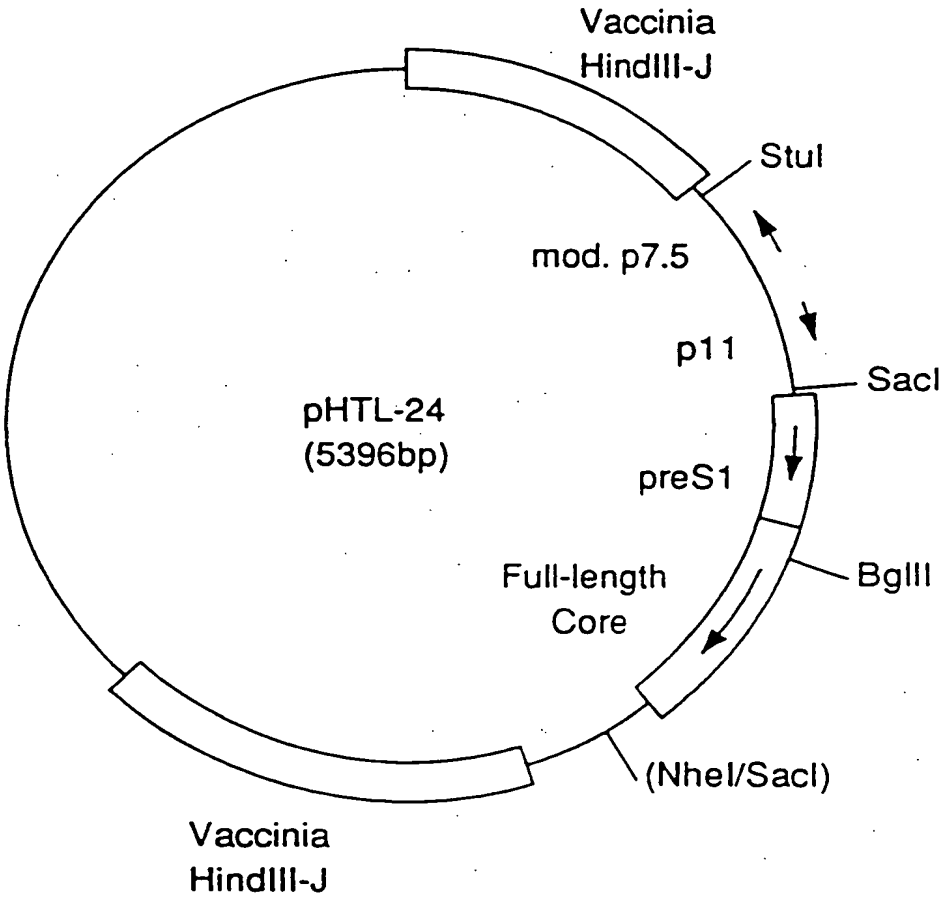

Vaccinia
HindIII-J
StuI
mod. p7.5
p11
SacI
pHTL-24
(5396bp)
preS1
Full-length
Core
BglII
(NheI/SacI)
Vaccinia
HindIII-J

FIG. 87

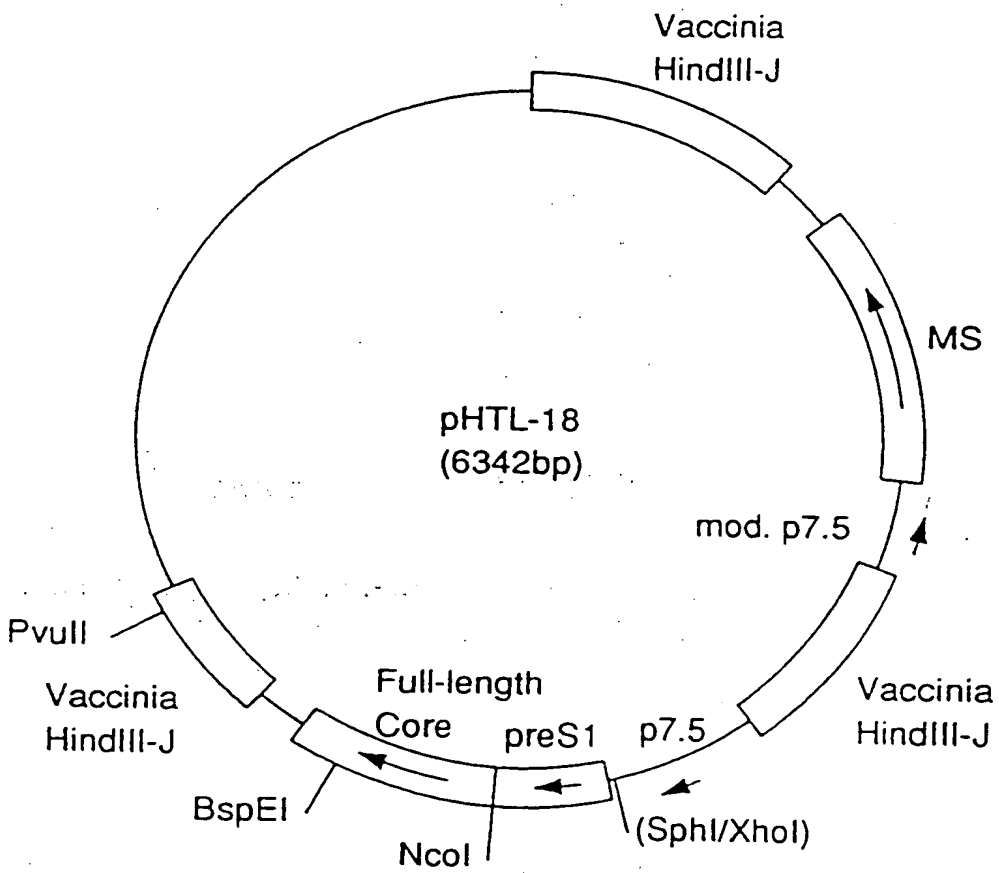

Vaccinia
HindIII-J
MS
pHTL-18
(6342bp)
mod. p7.5
PvuII
Vaccinia
HindIII-J
Full-length
Core
preS1
p7.5
Vaccinia
HindIII-J
BspEI
NcoI
(SphI/XhoI)

FIG. 88

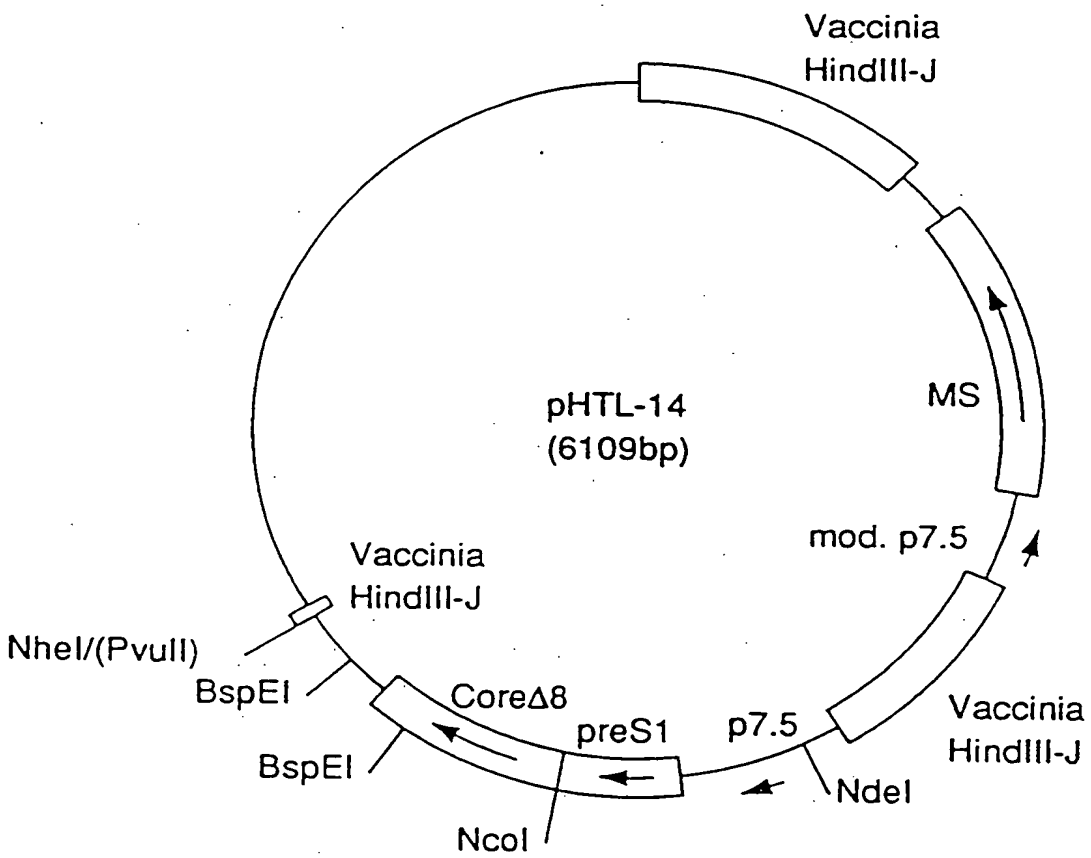

Vaccinia
HindIII-J
pHTL-14
(6109bp)
MS
mod. p7.5
Vaccinia
HindIII-J
Vaccinia
HindIII-J
NheI/(PvuII)
BspEI
BspEI
CoreΔ8
preS1
p7.5
NdeI
NcoI

FIG. 89

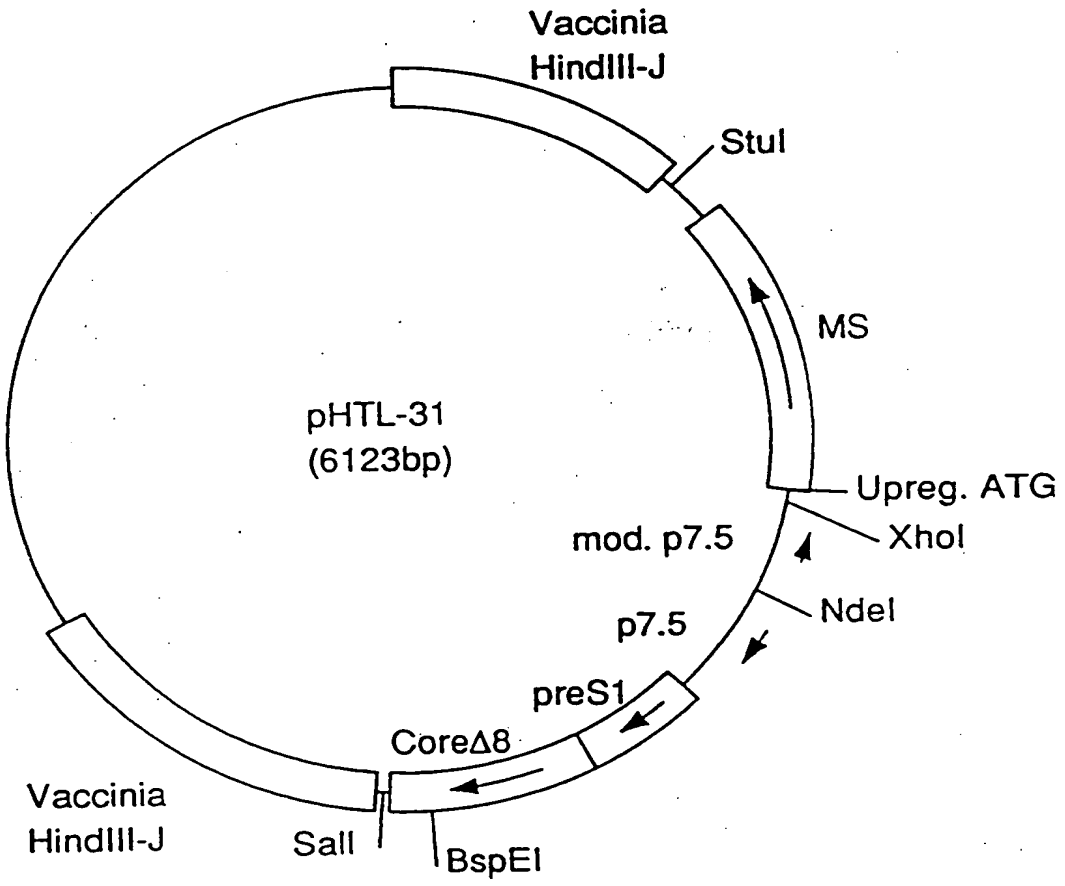
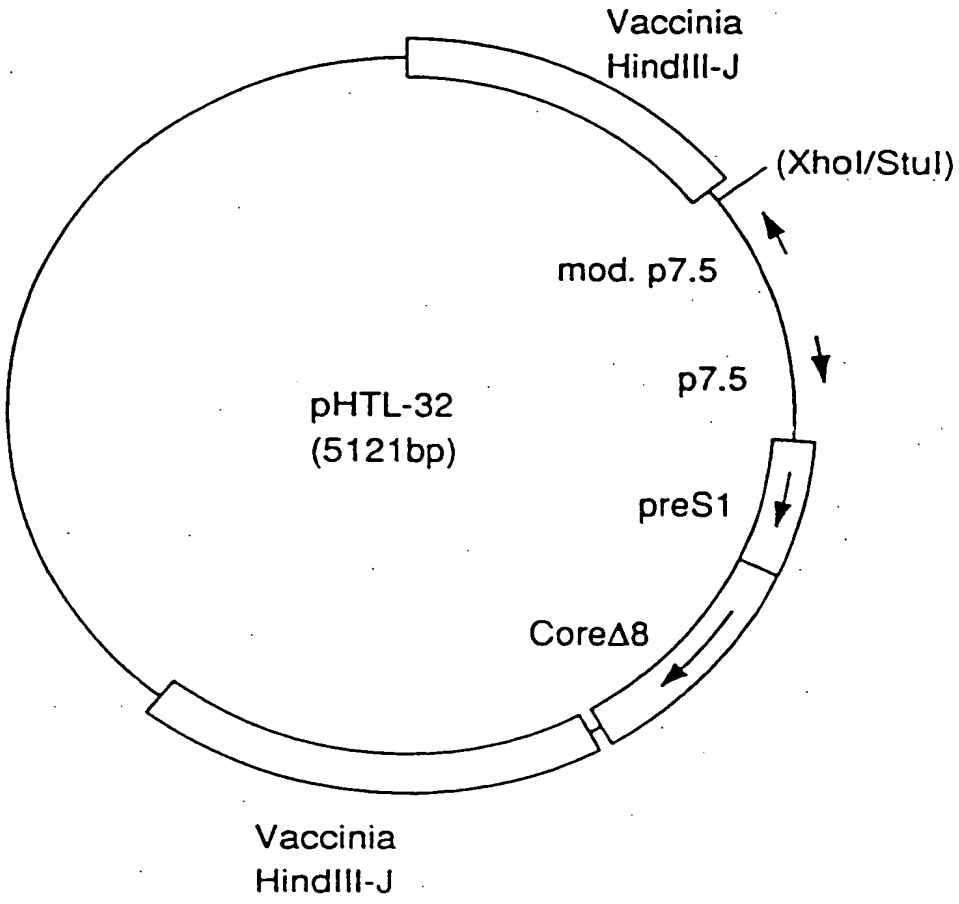

Vaccinia
HindIII-J
StuI
MS
pHTL-31
(6123bp)
Upreg. ATG
XhoI
mod. p7.5
NdeI
p7.5
preS1
CoreΔ8
Vaccinia
HindIII-J
SalI
BspEI

FIG. 90


Vaccinia
HindIII-J
(XhoI/StuI)
mod. p7.5
p7.5
pHTL-32
(5121bp)
preS1
CoreΔ8
Vaccinia
HindIII-J

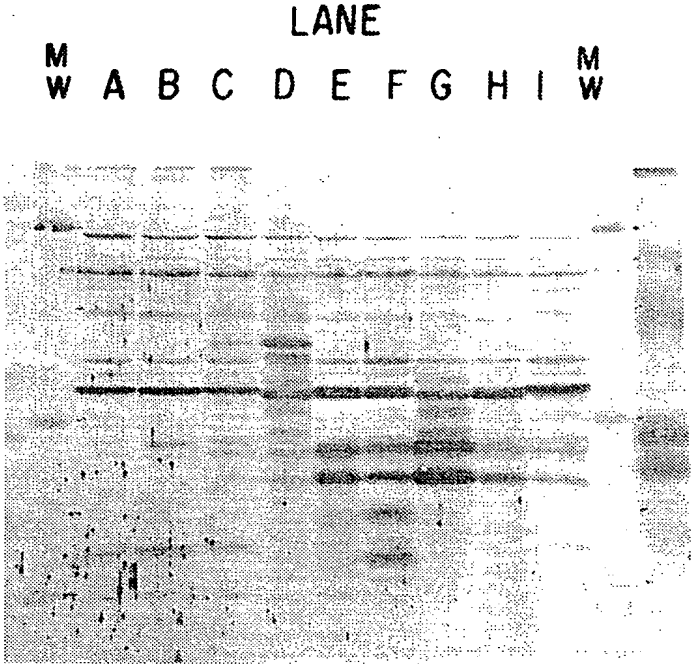


FIG. 91

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US93/11474

A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) :C12N 7/00, 7/01, 15/00, 15/11, 15/86; A16K 39/00, 39/29
US CL :435/69.1, 172.1, 172.3, 320.1; 424/93A; 530/350
According to International Patent Classification (IPC) or to both nati nal classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/69.1, 172.1, 172.3, 320.1; 424/93A; 530/350

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, Dialog, Medicine, Medline, Biotech
Search Terms: Vaccinia, Hepatitis, vaccine, core antigen

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Veterinary Parasitology, Volume 29, issued 1988, D.E. Hruby, "Present and Future Applications of Vaccinia Virus as a Vector," pages 281-292, See whole article, particularly Table 1 and Figure 3.	1-59
Y	Nature, Volume 311, issued August 1984, B. Moss et al., "Live Recombinant Vaccinia Virus Protects Chimpanzees Against Hepatitis B," pages 67-69, see whole article, particularly Figure 1 and 7th paragraph on page 68.	1-59

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be part of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search	Date of mailing of the international search report
17 JANUARY 1994	04 FEB 1994
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. NOT APPLICABLE	Auth rized fficer DAVID GUZO Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*